GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 12:20:54 ; Search time 2978.97 Seconds
(without alignments)
11596.087 Million cell updates/sec

Title: US-09-931-704-1
Perfect score: 797
Sequence: 1 attaaagcttcgccggagcc.....tctcctcttcgctccccccc 797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	797	100.0	797	6	AR002595	AR002595 Sequence
2	797	100.0	797	6	AX392086	AX392086 Sequence
3	797	100.0	797	6	BD195447	BD195447 The neuro
4	797	100.0	797	9	AF176911	AF176911 Homo sapi
5	793.4	99.5	881	6	AX205024	AX205024 Sequence
6	793.4	99.5	881	6	AX205042	AX205042 Sequence
7	776.4	97.4	1736	9	BC012939	BC012939 Homo sapi
8	749.4	94.0	1689	9	AF172854	AF172854 Homo sapi
9	749.4	94.0	1710	6	BD132824	BD132824 Cardiotro
10	680	85.3	680	9	AY049779	AY049779 Homo sapi
11	669.4	84.0	819	6	AR002597	AR002597 Sequence
12	669.4	84.0	819	6	AX392089	AX392089 Sequence
13	669.4	84.0	819	6	BD195449	BD195449 The neuro
14	669.4	84.0	819	10	AF176913	AF176913 Mus muscu
15	594.4	74.6	1692	6	AX205060	AX205060 Sequence
16	523.4	65.7	5087	6	AR002596	AR002596 Sequence
17	523.4	65.7	5087	6	AX392088	AX392088 Sequence
18	523.4	65.7	5087	6	BD195448	BD195448 The neuro
19	523.4	65.7	5087	9	AF176912	AF176912 Homo sapi
20	523.4	65.7	135116	2	AP002437	AP002437 Homo sapi
c 21	523.4	65.7	168565	9	AP003419	AP003419 Homo sapi
c 22	521.8	65.5	169144	2	AC005849	AC005849 Homo sapi
23	429.6	53.9	246093	2	AC140073	AC140073 Mus muscu
24	429.6	53.9	269155	10	AC109138	AC109138 Mus muscu
c 25	415.2	52.1	258710	2	AC135823	AC135823 Rattus no
26	269.4	33.8	283	6	AX202145	AX202145 Sequence
c 27	86.4	10.8	63347	2	AC110526	AC110526 Mus muscu
28	80.2	10.1	396	6	BD132825	BD132825 Cardiotro
c 29	68.2	8.6	233361	2	BX510356	BX510356 Danio rer
30	65	8.2	169873	2	BX324116	BX324116 Danio rer
c 31	59.4	7.5	125020	9	AF429315	AF429315 Homo sapi
32	57.8	7.3	125020	9	AF429315	AF429315 Homo sapi
33	52.2	6.5	900	10	AB125661	AB125661 Mus muscu
34	47.6	6.0	648	5	AF181480	AF181480 Plethodon
35	47.6	6.0	648	5	AF181481	AF181481 Plethodon
c 36	46.4	5.8	167672	2	AC132806	AC132806 Homo sapi
37	46.4	5.8	177719	9	AC135048	AC135048 Homo sapi
38	46.4	5.8	187727	2	AC021142	AC021142 Homo sapi
c 39	46.4	5.8	202414	2	AC135051	AC135051 Homo sapi
40	46.2	5.8	97826	9	AL591647	AL591647 Human DNA
c 41	46.2	5.8	173609	9	AL451077	AL451077 Human DNA
42	46	5.8	648	5	AF181482	AF181482 Plethodon
43	45.8	5.7	1107	6	BD180326	BD180326 Highly th
44	45.6	5.7	1901	5	GGLAMB1	X16878 Chicken mRN
45	44.8	5.6	16124	6	AX024277	AX024277 Sequence

ALIGNMENTS

RESULT 1
AR002595
LOCUS AR002595 797 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5741772.
ACCESSION AR002595
VERSION AR002595.1 GI:3964149
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 797)
AUTHORS Chang,M.-s.
TITLE Neurotrophic factor NNT-1
JOURNAL Patent: US 5741772-A 1 21-APR-1998;
FEATURES Location/Qualifiers

```
     source          1. .797
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            100.0%;  Score 797;  DB 6;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-152;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797


RESULT 2
AX392086
LOCUS       AX392086                797 bp    DNA     linear   PAT 23-MAR-2002
DEFINITION  Sequence 1 from Patent WO0215977.
ACCESSION   AX392086
VERSION     AX392086.1  GI:19700574
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Senaldi,G.
  TITLE     Methods and compositions for treating ige-related disease using nn
            t-1 inhibitors
  JOURNAL   Patent: WO 0215977-A 1 28-FEB-2002;
            Amgen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             90. .767
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28985.1"
                     /db_xref="GI:19700575"
                     /db_xref="REMTREMBL:CAD28985"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
     mat_peptide     171. .764
                     /product="Met at -27"
ORIGIN

  Query Match            100.0%;  Score 797;  DB 6;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-152;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
```

```
Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 3
BD195447
LOCUS BD195447 797 bp DNA linear PAT 17-JUL-2003
DEFINITION The neurotrophic factor NNT-1.
ACCESSION BD195447
VERSION BD195447.1 GI:33005217
KEYWORDS JP 2002514067-A/1.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 797)
AUTHORS Chang,M.S., Elliot,G.S., Senaldi,G. and Sarmiento,U.
TITLE The neurotrophic factor NNT-1
JOURNAL Patent: JP 2002514067-A 1 14-MAY-2002;
AMGEN INC
COMMENT OS Unidentified
PN JP 2002514067-A/1
PD 14-MAY-2002
PF 02-FEB-1998 JP 1998533258
PR 03-FEB-1997 US 08/792019,30-JAN-1998 US 09/016534 PI
MING SHI CHANG,GARY S ELLIOT,GIORGI SENALDI,ULLA SARMIENTO PC
C12N15/18,C07K14/475,C07K16/22,C12N1/21,A61K38/18,A01K67/027 CC
Strandedness: Single;
CC Topology: Linear;
CC The neurotrophic factor NNT-1
FH Key Location/Qualifiers
FT CDS 90. .764
FT mat peptide 171. .764
FT sig peptide 90. .170.
FEATURES Location/Qualifiers
source 1. .797
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 797; DB 6; Length 797;
Best Local Similarity 100.0%; Pred. No. 1.2e-152;
Matches 797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 4
AF176911
LOCUS AF176911 797 bp mRNA linear PRI 04-OCT-1999
DEFINITION Homo sapiens neurotrophin-1/B-cell stimulating factor-3 mRNA,
complete cds.
ACCESSION AF176911
VERSION AF176911.1 GI:6007640
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 797)
AUTHORS Senaldi,G., Varnum,B.C., Sarmiento,U., Starnes,C., Lile,J.,
Scully,S., Guo,J., Elliott,G., McNinch,J., Shaklee,C.L.,
Freeman,D., Manu,F., Simonet,W.S., Boone,T. and Chang,M.S.
TITLE Novel neurotrophin-1/B cell-stimulating factor-3: a cytokine of the
IL-6 family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11458-11463 (1999)
MEDLINE 99432254
PUBMED 10500198
REFERENCE 2 (bases 1 to 797)
AUTHORS Senaldi,G., Varnum,B., Sarmiento,U., Lile,J., Starnes,C.,
Scully,S., Guo,J., Elliott,G., McNinch,J., Freeman,D., Shaklee,C.,
Manu,F., Simonet,S., Boone,T. and Chang,M.-S.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1999) Amgen, Inc., One Amgen Center Drive,
Thousand Oaks, CA 91320, USA
FEATURES Location/Qualifiers
source 1. .797
/organism="Homo sapiens"

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             90. .767
                     /codon_start=1
                     /product="neurotrophin-1/B-cell stimulating factor-3"
                     /protein_id="AAF00991.1"
                     /db_xref="GI:6007641"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match            100.0%;  Score 797;  DB 9;  Length 797;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-152;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

```
RESULT 5
AX205024
LOCUS       AX205024                881 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 1 from Patent WO0155172.
ACCESSION   AX205024
VERSION     AX205024.1  GI:15394259
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elson,G., Gauchat,J.F., Plun-Favreau,H., Chevalier,S. and Gascan,H.
  TITLE     Isolated complex comprising a nnt-1 protein and in addition at
            least a clf-1 protein and/or a scntfr_g(a) protein
  JOURNAL   Patent: WO 0155172-A 1 02-AUG-2001;
            PIERRE FABRE MEDICAMENT (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE
            LA RECHERCHE MEDICALE (INSERM) (FR)
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             174. .851
                     /note="unnamed protein product; NNT-1"
                     /codon_start=1
                     /protein_id="CAC60175.1"
                     /db_xref="GI:15394260"
                     /db_xref="REMTREMBL:CAC60175"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match            99.5%;  Score 793.4;  DB 6;  Length 881;
  Best Local Similarity  99.9%;  Pred. No. 6.3e-152;
  Matches  794;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 144

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        145 CCGCACCCGGCCGGCCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 204

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 264

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 324

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 384

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 444

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 504

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 564
```

```
Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 624

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 684

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 744

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 804

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 864

Qy        781 CCTCTTCGCTCCCCC 795
              |||||||||||||||
Db        865 CCTCTTCGCTCCCCC 879
```

RESULT 6
AX205042
LOCUS AX205042 881 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0155219.
ACCESSION AX205042
VERSION AX205042.1 GI:15394277
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Elson,G. and Gauchat,J.F.
TITLE Scscntfr/nnt-1 fusion protein
JOURNAL Patent: WO 0155219-A 1 02-AUG-2001;
PIERRE FABRE MEDICAMENT (FR)
FEATURES Location/Qualifiers
source 1. .881
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 174. .851
/note="unnamed protein product; NNT-1"
/codon_start=1
/protein_id="CAC60178.1"
/db_xref="GI:15394278"
/db_xref="REMTREMBL:CAC60178"
/translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
HLGAHGF"
ORIGIN

Query Match 99.5%; Score 793.4; DB 6; Length 881;
Best Local Similarity 99.9%; Pred. No. 6.3e-152;
Matches 794; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 144

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 204

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 264

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 324

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 384

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 444

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 504

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 564

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 624

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 684

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 744

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 804

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 864

Qy        781 CCTCTTCGCTCCCCC 795
              |||||||||||||||
Db        865 CCTCTTCGCTCCCCC 879
```

RESULT 7
BC012939
LOCUS BC012939 1736 bp mRNA linear PRI 04-OCT-2003
DEFINITION Homo sapiens cardiotrophin-like cytokine, mRNA (cDNA clone
MGC:21195 IMAGE:4453813), complete cds.
ACCESSION BC012939
VERSION BC012939.1 GI:15277894
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1736)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

```
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1736)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 28 Row: b Column: 23
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 7019350.
FEATURES            Location/Qualifiers
     source         1. .1736
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="MGC:21195 IMAGE:4453813"
                    /tissue_type="Kidney, hypernephroma"
                    /clone_lib="NIH_MGC_89"
                    /lab_host="DH10B"
                    /note="Vector: pCMV-SPORT6"
     gene           1. .1736
                    /gene="CLC"
                    /note="synonym: BSF3"
                    /db_xref="LocusID:23529"
                    /db_xref="MIM:607672"
     CDS            73. .750
                    /codon_start=1
                    /product="cardiotrophin-like cytokine"
                    /protein_id="AAH12939.1"
                    /db_xref="GI:15277895"
                    /db_xref="LocusID:23529"
                    /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                    DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                    LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                    PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                    HLGAHGF"
     misc_feature   313. .693
                    /note="CNTF; Region: Ciliary neurotrophic factor"
                    /db_xref="CDD:pfam01110"
ORIGIN

  Query Match             97.4%;  Score 776.4;  DB 9;  Length 1736;
  Best Local Similarity   99.9%;  Pred. No. 1.7e-148;
  Matches  777;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         18 GCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCC 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCC 60

Qy         78 AGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGC 120

Qy        138 ACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCT 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCT 180

Qy        198 GGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGC 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGC 240

Qy        258 TTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCT 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCT 300

Qy        318 CCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGC 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGC 360

Qy        378 CTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTAC 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTAC 420

Qy        438 TTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTC 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTC 480

Qy        498 TGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTAC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTAC 540

Qy        558 CCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGT 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGT 600

Qy        618 GACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGG 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGG 660

Qy        678 CGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTC 720

Qy        738 ACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 778


RESULT 8
AF172854
LOCUS       AF172854               1689 bp    mRNA    linear   PRI 10-SEP-1999
DEFINITION  Homo sapiens cardiotrophin-like cytokine CLC mRNA, complete cds.
ACCESSION   AF172854
VERSION     AF172854.1  GI:5852980
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Shi,Y., Wang,W., Yourey,P.A., Gohari,S., Zukauskas,D., Zhang,J.,
            Ruben,S. and Alderson,R.F.
  TITLE     Computational EST database analysis identifies a novel member of
            the neuropoietic cytokine family
  JOURNAL   Biochem. Biophys. Res. Commun. 262 (1), 132-138 (1999)
  MEDLINE   99382254
   PUBMED   10448081
REFERENCE   2  (bases 1 to 1689)
```

```
  AUTHORS   Shi,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-1999) Molecular Biology, Human Genome Science,
            Inc., 9410 Key West Avenue, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .1689
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13.3"
     CDS             46. .723
                     /codon_start=1
                     /product="cardiotrophin-like cytokine CLC"
                     /protein_id="AAD54284.1"
                     /db_xref="GI:5852981"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match           94.0%;  Score 749.4;  DB 9;  Length 1689;
  Best Local Similarity  99.9%;  Pred. No. 5.5e-143;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCA 104
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCA 60

Qy        105 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 120

Qy        165 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 180

Qy        225 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 240

Qy        285 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 300

Qy        345 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 360

Qy        405 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 420

Qy        465 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 480

Qy        525 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 540

Qy        585 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 600

Qy        645 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 660

Qy        705 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 720

Qy        765 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||
Db        721 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 751


RESULT 9
BD132824
LOCUS       BD132824               1710 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Cardiotrophin-like cytokine.
ACCESSION   BD132824
VERSION     BD132824.1  GI:23227769
KEYWORDS    JP 2002507125-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1710)
  AUTHORS   Shi,Y. and Ruben,S.M.
  TITLE     Cardiotrophin-like cytokine
  JOURNAL   Patent: JP 2002507125-A 1 05-MAR-2002;
            HUMAN GENOME SCIENCES INC
COMMENT     PN   JP 2002507125-A/1
            PD   05-MAR-2002
            PF   29-JUN-1998 JP 1999505681
            PR   30-JUN-1997 US    60/051311
            PI   YANGGU SHI,STEVEN M RUBEN
            PC   C07K14/435,C07K14/52,C07K16/24,C12N1/21,C12N15/19,C12N15/64 CC
              Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   CDS             46. .720
            FT   sig peptide     46. .126
            FT   mat peptide     127. .720.
FEATURES             Location/Qualifiers
     source          1. .1710
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           94.0%;  Score 749.4;  DB 6;  Length 1710;
  Best Local Similarity  99.9%;  Pred. No. 5.5e-143;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCA 104
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCA 60

Qy        105 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 120

Qy        165 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 180

Qy        225 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 240

Qy        285 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 300

Qy        345 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 360

Qy        405 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 420
```

```
Qy        465 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 480

Qy        525 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 540

Qy        585 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 600

Qy        645 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 660

Qy        705 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 720

Qy        765 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||
Db        721 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 751
```

RESULT 10
AY049779
LOCUS AY049779 680 bp mRNA linear PRI 23-OCT-2001
DEFINITION Homo sapiens cardiotrophin-like cytokine (CLC) mRNA, complete cds.
ACCESSION AY049779
VERSION AY049779.1 GI:16356642
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Hu,X., Xu,Y., Zhang,B., Peng,X., Yuan,J. and Qiang,B.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2001) Department of Biochemistry, Institute of Basic Medical Science, Chinese Academy of Medical Sciences, 5 Dong Dan San Tiao, Beijing 100005, P.R. China
FEATURES Location/Qualifiers
source 1. .680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene 1. .680
/gene="CLC"
CDS 3. .680
/gene="CLC"
/note="neurotrophin-1/B-cell stimulating factor-3"
/codon_start=1
/product="cardiotrophin-like cytokine"
/protein_id="AAL15436.1"
/db_xref="GI:16356643"
/translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
HLGAHGF"
ORIGIN

Query Match 85.3%; Score 680; DB 9; Length 680;
Best Local Similarity 100.0%; Pred. No. 9.5e-129;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         88 CCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCT 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCT 60

Qy        148 GGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCA 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCA 120

Qy        208 TCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGA 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGA 180

Qy        268 CCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 240

Qy        328 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 300

Qy        388 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 360

Qy        448 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 420

Qy        508 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 567
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 480

Qy        568 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 540

Qy        628 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 600

Qy        688 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 660

Qy        748 TGGGGGCTCATGGCTTCTGA 767
              ||||||||||||||||||||
Db        661 TGGGGGCTCATGGCTTCTGA 680
```

RESULT 11
AR002597
LOCUS AR002597 819 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 4 from patent US 5741772.
ACCESSION AR002597
VERSION AR002597.1 GI:3964151
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 819)
AUTHORS Chang,M.-s.
TITLE Neurotrophic factor NNT-1
JOURNAL Patent: US 5741772-A 4 21-APR-1998;
FEATURES Location/Qualifiers
source 1. .819
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 84.0%; Score 669.4; DB 6; Length 819;
Best Local Similarity 92.0%; Pred. No. 1.3e-126;
Matches 717; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
```

```
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              ||||||||| ||||| |||||||||||| |||||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| |||||||||||||||| ||   ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||  |||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| ||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 12
AX392089
LOCUS       AX392089                819 bp    DNA     linear   PAT 23-MAR-2002
DEFINITION  Sequence 4 from Patent WO0215977.
ACCESSION   AX392089
VERSION     AX392089.1  GI:19700577
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Senaldi,G.
  TITLE     Methods and compositions for treating ige-related disease using nn
            t-1 inhibitors
  JOURNAL   Patent: WO 0215977-A 4 28-FEB-2002;
            Amgen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Mus sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10095"
     CDS             95. .772
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28986.1"
                     /db_xref="GI:19700578"
                     /db_xref="REMTREMBL:CAD28986"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQ
                     PLPGTEPAWAPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTL
                     HLEAHGF"
     sig_peptide     95. .175
     mat_peptide     176. .769
                     /product="unnamed"
ORIGIN

  Query Match           84.0%;  Score 669.4;  DB 6;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.3e-126;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              ||||||| ||||| |||||||||| |||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| |||||||||||||||| ||   ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||  |||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| ||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 13
BD195449
LOCUS       BD195449                819 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  The neurotrophic factor NNT-1.
```

```
ACCESSION   BD195449
VERSION     BD195449.1  GI:33005219
KEYWORDS    JP 2002514067-A/3.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Chang,M.S., Elliot,G.S., Senaldi,G. and Sarmiento,U.
  TITLE     The neurotrophic factor NNT-1
  JOURNAL   Patent: JP 2002514067-A 3 14-MAY-2002;
            AMGEN INC
COMMENT     OS   Unidentified
            PN   JP 2002514067-A/3
            PD   14-MAY-2002
            PF   02-FEB-1998 JP 1998533258
            PR   03-FEB-1997 US    08/792019,30-JAN-1998 US    09/016534 PI
            MING SHI CHANG,GARY S ELLIOT,GIORGI SENALDI,ULLA SARMIENTO PC
            C12N15/18,C07K14/475,C07K16/22,C12N1/21,A61K38/18,A01K67/027 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   The neurotrophic factor NNT-1
            FH   Key             Location/Qualifiers
            FT   CDS             95. .769
            FT   mat peptide     176. .769
            FT   sig peptide     95. .175.
FEATURES                Location/Qualifiers
     source             1. .819
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
ORIGIN

  Query Match             84.0%;  Score 669.4;  DB 6;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.3e-126;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| ||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || |||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| |||| || |||||||||| |||||||||||||||| || ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||  ||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| |||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              |||||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| |||||||| |||||||
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 14
AF176913
LOCUS       AF176913                819 bp    mRNA    linear   ROD 04-OCT-1999
DEFINITION  Mus musculus neurotrophin-1/B-cell stimulating factor-3 mRNA,
            complete cds.
ACCESSION   AF176913
VERSION     AF176913.1  GI:6007644
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Senaldi,G., Varnum,B.C., Sarmiento,U., Starnes,C., Lile,J.,
            Scully,S., Guo,J., Elliott,G., McNinch,J., Shaklee,C.L.,
            Freeman,D., Manu,F., Simonet,W.S., Boone,T. and Chang,M.S.
  TITLE     Novel neurotrophin-1/B cell-stimulating factor-3: a cytokine of the
            IL-6 family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11458-11463 (1999)
  MEDLINE   99432254
   PUBMED   10500198
REFERENCE   2  (bases 1 to 819)
  AUTHORS   Senaldi,G., Varnum,B., Sarmiento,U., Lile,J., Starnes,C.,
            Scully,S., Guo,J., Elliott,G., McNinch,J., Freeman,D., Shaklee,C.,
            Manu,F., Simonet,S., Boone,T. and Chang,M.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1999) Amgen, Inc., One Amgen Center Drive,
            Thousand Oaks, CA 91320, USA
FEATURES                Location/Qualifiers
     source             1. .819
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /db_xref="taxon:10090"
     CDS                95. .772
                        /codon_start=1
                        /product="neurotrophin-1/B-cell stimulating factor-3"
                        /protein_id="AAF00993.1"
                        /db_xref="GI:6007645"
                        /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                        DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLR
                        LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQ
                        PLPGTEPAWAPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTL
                        HLEAHGF"
ORIGIN

  Query Match             84.0%;  Score 669.4;  DB 10;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.3e-126;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124
```

```
Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              ||||||||  |||||  ||||||||||| ||||||||||||||||||||||||||||||  ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              |||||  ||  |||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              |||||  ||||||||||||||  ||||||||||  ||||||||||||||  ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              |||||  ||||||||  |||||  ||  ||||||||||  ||||||||||||||||||  ||  ||
Db        305 GAGCCTGACTTCAATCCTCCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              |||||  ||||||||||||||||||||  ||||||||||||||||||||  |||||  |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              ||  |||||  ||||||||||||||||||||||||||||||||||||  |||||  ||  ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              ||  ||||||||||||||||  ||||||||||||||||||||||||||||||||||  |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  |  ||  ||||||||||||||||||||  |||||  |||||||  ||  | |||  |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||  |||||  ||  ||||||||||||||||||  ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              ||||||||||||  ||||||||||||||  |||  |||  |||||  |||||||  |||||||  |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783
```

RESULT 15
AX205060
LOCUS AX205060 1692 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 19 from Patent WO0155219.
ACCESSION AX205060
VERSION AX205060.1 GI:15394299
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Elson,G. and Gauchat,J.F.
TITLE Scscntfr/nnt-1 fusion protein
JOURNAL Patent: WO 0155219-A 19 02-AUG-2001;
PIERRE FABRE MEDICAMENT (FR)
FEATURES Location/Qualifiers
source 1. .1692
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
CDS 1. .1692
/note="unnamed protein product; proteine de fusion"
/codon_start=1
/transl_table=11
/protein_id="CAC60181.1"
/db_xref="GI:15394300"
/db_xref="REMTREMBL:CAC60181"
/translation="MAAPVPWACCAVLAAAAAVVYAQRHSPQEAPHVQYERLGSDVTL

```
                  PCGTANWDAAVTWRVNGTDLAPDLLNGSQLVLHGLELGHSGLYACFHRDSWHLRHQVL
                  LHVGLPPREPVLSCRSNTYPKGFYCSWHLPTPTYIPNTFNVTVLHGSKIMVCEKDPAL
                  KNRCHIRYMHLFSTIKYKVSISVSNALGHNATAITFDEPTIVKPDPPENVVARPVPSN
                  PRRLEVTWQTPSTWPDPESFPLKFFLRYRPLILDQWQHVELSDGTAHTITDAYAGKEY
                  IIQVAAKDNEIGTWSDWSVAAHATPWTEEPRHLTTEAQAAETTTSTTSSLAPPPTTKI
                  CDLEGGGGSGGGGSGGGGSLELNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNY
                  LGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR
                  QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ
                  KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGFEQKLISEEDL"
ORIGIN

Query Match 74.6%; Score 594.4; DB 6; Length 1692;
Best Local Similarity 98.2%; Pred. No. 2.5e-111;
Matches 601; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy        171 CTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACC 230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1066 CTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACC 1125

Qy        231 CGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCC 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 CGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCC 1185

Qy        291 CCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCC 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 CCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCC 1245

Qy        351 ACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTAC 410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 ACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTAC 1305

Qy        411 GAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCT 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 GAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCT 1365

Qy        471 GAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATT 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 GAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATT 1425

Qy        531 GCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCC 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1426 GCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCC 1485

Qy        591 ACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTG 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1486 ACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTG 1545

Qy        651 CTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1546 CTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAG 1605

Qy        711 AAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTT 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1606 AAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCGAGCAG 1665

Qy        771 CTGACCTTCTCC 782
                 |    |||||
Db       1666 AAGCTGATCTCC 1677
```

Search completed: August 15, 2004, 23:46:54
Job time : 2983.97 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 15:26:00 ; Search time 1814.21 Seconds
(without alignments)
13118.770 Million cell updates/sec

Title: US-09-931-704-1
Perfect score: 797
Sequence: 1 attaaagcttcgccggagcc.....tctcctcttcgctccccccc 797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	740.4	92.9	887	13	BQ948158	BQ948158 AGENCOURT
	2	711.6	89.3	788	14	CF271927	CF271927 AGENCOURT
	3	706.8	88.7	787	14	CK130210	CK130210 AGENCOURT
	4	700	87.8	806	14	CK130211	CK130211 AGENCOURT
	5	698.6	87.7	804	14	CF271929	CF271929 AGENCOURT
	6	694.8	87.2	752	14	CF271925	CF271925 AGENCOURT
	7	684	85.8	810	14	CK130212	CK130212 AGENCOURT
	8	662	83.1	662	29	AY409170	AY409170 Homo sapi
	9	656.4	82.4	1157	13	BQ940483	BQ940483 AGENCOURT
	10	654	82.1	662	29	AY409171	AY409171 Pan trogl
c	11	647.2	81.2	800	14	CF271926	CF271926 AGENCOURT
	12	641.6	80.5	853	12	BG437538	BG437538 602489277
	13	621	77.9	1053	12	BG164929	BG164929 602343555
c	14	604.8	75.9	801	14	CF271930	CF271930 AGENCOURT
c	15	602	75.5	763	14	CF271924	CF271924 AGENCOURT
	16	590.8	74.1	594	12	BM763333	BM763333 K-EST0044
	17	577.4	72.4	580	12	BM848189	BM848189 K-EST0127
	18	576.4	72.3	678	13	BY735168	BY735168 BY735168
	19	572.4	71.8	662	29	AY409172	AY409172 Mus muscu
	20	571.4	71.7	573	12	BM840863	BM840863 K-EST0118
	21	571.2	71.7	1028	9	AL543945	AL543945 AL543945
	22	549.8	69.0	634	13	BX497225	BX497225 DKFZp779H
c	23	539.8	67.7	793	14	CK032835	CK032835 AGENCOURT
	24	528.8	66.3	532	12	BM846370	BM846370 K-EST0125
	25	520.4	65.3	522	12	BM841897	BM841897 K-EST0119
	26	518.4	65.0	542	12	BM821005	BM821005 K-EST0089
	27	514.8	64.6	529	12	BM847924	BM847924 K-EST0127
	28	458.8	57.6	569	12	BM846748	BM846748 K-EST0125
	29	439.6	55.2	831	28	CC138171	CC138171 NDL.3C1.T
	30	412	51.7	476	10	BE632644	BE632644 uv63f01.y
	31	405.4	50.9	420	12	BM764238	BM764238 K-EST0045
	32	403.4	50.6	458	12	BM363136	BM363136 BS320054A
	33	387	48.6	407	9	AI752561	AI752561 cn17d05.x
	34	379.8	47.7	448	13	BY531444	BY531444 BY531444
	35	371.4	46.6	482	12	BM846622	BM846622 K-EST0125
c	36	370.2	46.4	766	14	CF271928	CF271928 AGENCOURT
	37	331.2	41.6	505	13	BX528877	BX528877 BX528877
	38	329.6	41.4	388	13	BY022718	BY022718 BY022718
	39	319.8	40.1	668	13	BY734701	BY734701 BY734701
	40	301.2	37.8	354	13	BY015239	BY015239 BY015239
	41	277.6	34.8	913	12	BI912197	BI912197 603070588
	42	274.4	34.4	512	9	AI390475	AI390475 mu59g08.y
	43	258.4	32.4	392	13	BY012378	BY012378 BY012378
	44	242.6	30.4	663	13	BY722568	BY722568 BY722568
	45	227.2	28.5	488	9	AA015243	AA015243 mh30c11.r

ALIGNMENTS

RESULT 1
BQ948158
LOCUS BQ948158 887 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8813192 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6428214 5', mRNA sequence.
ACCESSION BQ948158
VERSION BQ948158.1 GI:22363636
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2614 row: h column: 07
High quality sequence stop: 674.

```
FEATURES             Location/Qualifiers
     source          1. .887
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6428214"
                     /tissue_type="epidermoid carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_101"
                     /note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
                     XhoI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             92.9%;  Score 740.4;  DB 13;  Length 887;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-143;
  Matches  752;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 67

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         68 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 127

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 187

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 247

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 307

Qy        301 AGCCAGACTTCAACCCTCCCCGCC-TGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        308 AGCCAGACTTCAACCCTCCCCGCCATGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 367

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 427

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 487

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 547

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 607

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 667

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 727

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGG 753
              |||||||||||||||||||||||||||||||||
Db        728 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGG 761


RESULT 2
CF271927
LOCUS       CF271927               788 bp    mRNA    linear   EST 26-NOV-2003
DEFINITION  AGENCOURT_15196974 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002180 5', mRNA sequence.
ACCESSION   CF271927
VERSION     CF271927.2  GI:38558834
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627839.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 10
            High quality sequence start: 12
            High quality sequence stop: 761.
FEATURES             Location/Qualifiers
     source          1. .788
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002180"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             89.3%;  Score 711.6;  DB 14;  Length 788;
  Best Local Similarity   99.4%;  Pred. No. 6e-137;
  Matches  714;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         77 CAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTG 136
              ||| |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 CAGTCGACGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTG 83

Qy        137 CACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCC 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         84 CACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCC 143

Qy        197 TGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAG 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAG 203

Qy        257 CTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 CTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCC 263

Qy        317 TCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 TCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAG 323

Qy        377 CCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 CCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTA 383

Qy        437 CTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTT 443

Qy        497 CTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTA 503

Qy        557 CCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAG 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        504 CCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAG 563

Qy        617 TGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTG 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        564 TGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTG 623

Qy        677 GCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        624 GCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGT 683

Qy        737 CACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCC 794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        684 CACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCC 741
```

RESULT 3
CK130210

```
LOCUS       CK130210                787 bp    mRNA    linear   EST 02-DEC-2003
DEFINITION  AGENCOURT_15196910 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002178 5', mRNA sequence.
ACCESSION   CK130210
VERSION     CK130210.1  GI:38622146
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 08
            High quality sequence stop: 715.
```

```
FEATURES             Location/Qualifiers
     source          1. .787
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002178"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 88.7%; Score 706.8; DB 14; Length 787;
Best Local Similarity 99.4%; Pred. No. 6e-136;
Matches 708; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy         84 AGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTG 143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTG 60

Qy        144 CTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCC 203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTGGCACCTCCCTGCAGTGCCAGCTGTCAATCGCACAGGGGACCCAGGGCCTGGCCCC 120

Qy        204 TCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCT 263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCT 180

Qy        264 GGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGC 323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGC 240

Qy        324 CTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAAT 383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAAT 300

Qy        384 GACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGT 443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGT 360

Qy        444 GGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACC 503
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCCTCAACCGTCGGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACC 420

Qy        504 AGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTG 563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTG 480

Qy        564 CCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTC 623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTC 540

Qy        624 CTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCG 683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCG 600

Qy        684 GCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTG 743
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTG 660

Qy        744 CACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
                ||||||||||||||||||||||||||||||||||||||||||||||||||   |
Db        661 CACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCNNC 712


RESULT 4
CK130211
LOCUS       CK130211                806 bp    mRNA    linear   EST 02-DEC-2003
DEFINITION  AGENCOURT_15196878 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002177 5', mRNA sequence.
ACCESSION   CK130211
VERSION     CK130211.1  GI:38622147
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 07
            High quality sequence start: 11
            High quality sequence stop: 736.
FEATURES             Location/Qualifiers
     source          1. .806
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002177"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             87.8%;  Score 700;  DB 14;  Length 806;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-134;
  Matches  700;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          96 CTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTC 155
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         38 CTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTC 97

Qy        156 CCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAA 215
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 CCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAA 157

Qy        216 ACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTG 275
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 ACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTG 217

Qy        276 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAG 335
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAG 277

Qy        336 ACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGG 395
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 ACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGG 337

Qy        396 CTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGT 455
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 CTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGT 397

Qy        456 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 515
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 457

Qy        516 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 575
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 517

Qy        576 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 635
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 577

Qy        636 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTC 695
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTC 637

Qy        696 AACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCT 755
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 AACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCT 697

Qy        756 CATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
               ||||||||||||||||||||||||||||||||||||||||
Db        698 CATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 737


RESULT 5
CF271929
LOCUS       CF271929                804 bp    mRNA    linear   EST 26-NOV-2003
DEFINITION  AGENCOURT_15196942 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002179 5', mRNA sequence.
ACCESSION   CF271929
VERSION     CF271929.2  GI:38558835
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 804)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627841.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
```

```
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 09
            High quality sequence start: 13
            High quality sequence stop: 711.
FEATURES             Location/Qualifiers
     source          1. .804
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002179"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            87.7%;  Score 698.6;  DB 14;  Length 804;
  Best Local Similarity  99.2%;  Pred. No. 3e-134;
  Matches  712;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy         77 CAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTG 136
              |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 CAGTCGACGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTG 84

Qy        137 CACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCC 196
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 CACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCC 144

Qy        197 TGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAG 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAG 204

Qy        257 CTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 CTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCC 264

Qy        317 TCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAG 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAG 324

Qy        377 CCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 CCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTA 384

Qy        437 CTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTT 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTT 444

Qy        497 CTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTA 504

Qy        557 CCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAG 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAG 564

Qy        617 TGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTG 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 TGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTG 624

Qy        677 GCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 GCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGT 684

Qy        737 CACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCC 794
              |||||||||||||||||||||||||||||||||||| |||||||||||||||| |
Db        685 CACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGA-CTTCTCCTCTTCGCTCCNC 741


RESULT 6
CF271925
LOCUS       CF271925                752 bp    mRNA    linear   EST 26-NOV-2003
DEFINITION  AGENCOURT_15197006 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002181 5', mRNA sequence.
ACCESSION   CF271925
VERSION     CF271925.2  GI:38558833
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 752)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627837.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 11
            High quality sequence stop: 743.
FEATURES             Location/Qualifiers
     source          1. .752
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002181"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN
```

```
  Query Match             87.2%;  Score 694.8;  DB 14;  Length 752;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-133;
  Matches  707;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy         85 GCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGC 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGC 65

Qy        145 TCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCT 204
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 TCTGGCACCTCCTTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCT 125

Qy        205 CCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTG 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTG 185

Qy        265 GGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCC 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCC 245

Qy        325 TGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATG 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATG 305

Qy        385 ACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTG 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTG 365

Qy        445 GCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCA 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCA 425

Qy        505 GCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGC 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGC 485

Qy        565 CCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCC 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCC 545

Qy        625 TCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGG 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 TCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGG 605

Qy        685 CCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGC 744
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        606 CCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCCGC 665

Qy        745 ACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCC 794
              |||||||||||||||||||||||||||||||| ||||||||||||
Db        666 ACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT-CTCTTCGCTCCCC 714


RESULT 7
CK130212
LOCUS       CK130212                 810 bp    mRNA    linear   EST 02-DEC-2003
DEFINITION  AGENCOURT_15196846 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002176 5', mRNA sequence.
ACCESSION   CK130212
VERSION     CK130212.1  GI:38622148
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 810)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
            cDNA Library Preparation: Bhat Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 06
            High quality sequence start: 13
            High quality sequence stop: 731.
FEATURES             Location/Qualifiers
     source          1. .810
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002176"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             85.8%;  Score 684;  DB 14;  Length 810;
  Best Local Similarity   99.3%;  Pred. No. 3.2e-131;
  Matches  687;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         78 AGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGC 137
              || |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 AGTCGACGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGC 84

Qy        138 ACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCT 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCT 144

Qy        198 GGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGC 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGC 204

Qy        258 TTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCT 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCT 264

Qy        318 CCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGC 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGC 324

Qy        378 CTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTAC 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 CTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTAC 384

Qy        438 TTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTC 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 TTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTC 444

Qy        498 TGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTAC 557
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        445 TGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTAC 504

Qy        558 CCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGT 617
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGT 564

Qy        618 GACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGG 677
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGG 624

Qy        678 CGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTC 737
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 CGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTC 684

Qy        738 ACCCTGCACCTGGGGGCTCATGGCTTCTGACT 769
              |||||||||||||||||||||||||||||| |
Db        685 ACCCTGCACCTGGGGGCTCATGGCTTCTGATT 716


RESULT 8
AY409170
LOCUS       AY409170                662 bp    DNA     linear   GSS 12-DEC-2003
DEFINITION  Homo sapiens HCM3468 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY409170
VERSION     AY409170.1  GI:39765138
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 662)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 662)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .662
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene            <1. .>662
                     /locus_tag="HCM3468"
ORIGIN

  Query Match            83.1%;  Score 662;  DB 29;  Length 662;
  Best Local Similarity  100.0%;  Pred. No. 1e-126;
  Matches  662;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        106 GGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGC 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGC 60

Qy        166 CAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACC 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACC 120

Qy        226 TCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGG 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGG 180

Qy        286 GCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCA 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCA 240

Qy        346 GGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGA 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGA 300

Qy        406 ACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCA 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCA 360

Qy        466 CTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCA 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCA 420

Qy        526 GCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTG 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTG 480

Qy        586 AACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCT 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCT 540

Qy        646 GGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCA 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCA 600

Qy        706 AGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCT 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCT 660

Qy        766 GA 767
              ||
Db        661 GA 662


RESULT 9
BQ940483
LOCUS       BQ940483                1157 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8864294 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6197786 5', mRNA sequence.
ACCESSION   BQ940483
VERSION     BQ940483.1  GI:22355961
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1157)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13608  row: g  column: 03
            High quality sequence stop: 572.
FEATURES             Location/Qualifiers
     source          1. .1157
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6197786"
                     /sex="male"
```

/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sciatic_nerve"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCCG-3' and 5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 82.4%; Score 656.4; DB 13; Length 1157;
Best Local Similarity 97.4%; Pred. No. 1.9e-125;
Matches 678; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

```
Qy      104 AGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGT 163
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       58 AGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGT 117

Qy      164 GCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGA 223
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      118 GCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGA 177

Qy      224 CCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCT 283
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      178 CCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCT 237

Qy      284 GGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCC 343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      238 GGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCC 297

Qy      344 CAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCA 403
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      298 CAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCA 357

Qy      404 GAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGC 463
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      358 GAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGC 417

Qy      464 CACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGG 523
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 CACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGG 477

Qy      524 CAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGAC 583
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 CAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGAC 537

Qy      584 TGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTT 643
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 TGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTT 597

Qy      644 CTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCT 703
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      598 CTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCT 657

Qy      704 CAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGG--CTCATGGC 761
            ||||||||||||||||||||||||||||||||||||||||||||||||||     |||||
Db      658 CAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGGCTCATGGGC 717

Qy      762 TTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCCC 797
            ||||||||||||||| | |  | | | ||||||
Db      718 TTCTGACTTCTGACCCTTCTCCTCTTTCGCTCCCCC 753
```

RESULT 10
AY409171
LOCUS AY409171 662 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM3468 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY409171
VERSION AY409171.1 GI:39765139
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 662)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 662)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1. .662
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene <1. .>662
/locus_tag="HCM3468"

ORIGIN

Query Match 82.1%; Score 654; DB 29; Length 662;
Best Local Similarity 99.2%; Pred. No. 4.7e-125;
Matches 657; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      106 GGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGC 165
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        1 GGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCATCTCCCTGCAGTGC 60

Qy      166 CAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACC 225
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACC 120

Qy      226 TCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGG 285
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      121 TCACCCGCTACCTGGAGCATCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGG 180

Qy      286 GCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCA 345
            |||||||||||||||||| |||||||||||||||||| ||||||||||||||||||||||
Db      181 GCCCCCCTTTCAACGAGCCTGACTTCAACCCTCCCCGTCTGGGGGCAGAGACTCTGCCCA 240

Qy      346 GGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGA 405
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      241 GGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGAGGCTGACCCAGA 300

Qy      406 ACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCA 465
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCA 360

Qy      466 CTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCA 525
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCA 420

Qy      526 GCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTG 585
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTG 480

Qy      586 AACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCT 645
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCT 540

Qy        646 GGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCA 705
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCA 600

Qy        706 AGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCT 765
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCT 660

Qy        766 GA 767
                ||
Db        661 GA 662


RESULT 11
CF271926/c
LOCUS       CF271926                800 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AGENCOURT_15196988 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002180 5', mRNA sequence.
ACCESSION   CF271926
VERSION     CF271926.2  GI:38453379
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627838.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 10
            High quality sequence start: 8
            High quality sequence stop: 720.
FEATURES              Location/Qualifiers
     source           1. .800
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:7002180"
                      /tissue_type="mixed"
                      /lab_host="DH5A (T1 phage-resistant)"
                      /clone_lib="NIH_MGC_195"
                      /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                      loxP-HindIII; Clones from this library have been
                      PCR-amplified using gene-specific primers to contain the
                      complete open reading frame (based on known gene sequences
                      available from NCBI's RefSeq). Template for PCR is cDNA
                      derived from either pooled cytoplasmic polyA RNA from 30
                      cells lines or pooled total RNA from 10 different tissues
                      (from BD Biosciences/Clontech and Washington University).
                      PCR products are directionally cloned into the loxP sites
                      of the pDNR-Dual vector. Library constructed by Dr.
                      Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                      Expression Laboratory, Research Technology Program, SAIC
                      Frederick, NCI-Frederick, Frederick, MD 21702). For
                      information on which gene each clone represents, please
                      visit our anonymous ftp site at
                      ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                      a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           81.2%;  Score 647.2;  DB 14;  Length 800;
  Best Local Similarity   99.2%;  Pred. No. 1.3e-123;
  Matches  660;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        131 CCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCC 190
                ||  |||||||| ||| ||||||||||||||||||| ||||||||||||||||||||
Db        753 CCCTTGCACGGTGTTCT-GCACCTCCCTGCAGTGCCAGCTNTCAATCGCACAGGGGACCC 695

Qy        191 AGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACT 250
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 AGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACT 635

Qy        251 CCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTT 310
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 CCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTT 575

Qy        311 CAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTG 370
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTG 515

Qy        371 GCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCT 430
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 GCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCT 455

Qy        431 GTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGC 490
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 GTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGC 395

Qy        491 CCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCT 550
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 CCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCT 335

Qy        551 GGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGC 610
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 GGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGC 275

Qy        611 CCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTG 670
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 CCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTG 215

Qy        671 GCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGC 730
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 GCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGC 155

Qy        731 TGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCT 790
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 TGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCT 95

Qy        791 CCCCC 795
                |||||
Db         94 CCCCC 90


RESULT 12
BG437538
LOCUS       BG437538                853 bp    mRNA    linear   EST 14-MAR-2001
DEFINITION  602489277F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4621530 5',
            mRNA sequence.
ACCESSION   BG437538
VERSION     BG437538.1  GI:13344044
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 853)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1384  row: i  column: 19
            High quality sequence stop: 761.
FEATURES             Location/Qualifiers
     source          1. .853
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4621530"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_18"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             80.5%;  Score 641.6;  DB 12;  Length 853;
  Best Local Similarity   96.9%;  Pred. No. 2e-122;
  Matches  729;  Conservative    0;  Mismatches   14;  Indels    9;  Gaps    7;

Qy         52 GGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACT 111
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db          1 GGAGAGGAGCCGCACCCGGCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACT 60

Qy        112 CGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTC 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTC 120

Qy        172 TCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCC 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCC 180

Qy        232 GCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCC 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCC 240

Qy        292 CTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCA 300

Qy        352 CTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACG 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACG 360

Qy        412 AGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTG 471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTG 420

Qy        472 AGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTG 531
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGCTGCGCCGCA-CCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTG 479

Qy        532 CGGGCGTCATGGCAGCTCTGGGCTACCCAC-TGCCCCAGCCGCTGCCTGGGAC-TGAACC 589
              |||||||||||||||||||||||||||||| || |||||||||||||||||| ||||||
Db        480 CGGGCGTCATGGCAGCTCTGGGCTACCCACTTGACCCAGCCGCTGCCTGGGACTTGAACC 539

Qy        590 CACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCT 649
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        540 CACTTGGACTCCTGGCCTTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCT 599

Qy        650 GCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAA-CCGGCTCAAGA 708
              |||||||||||||| |||||||||||||||||||||||||| ||||||||| ||||||||||
Db        600 GCTGAAGGAGCTGAAGACCTGGCTGTGGCGCTCGGCCACGGACTTCAAGCCGGCTCAAGA 659

Qy        709 AGAAGATGCAGCCTCCAG-CAGCTGCAGTCACCC---TGCACCTGGGGGCTCATGGCTTC 764
              |||||||||||||||||| || |||||||||| ||        | ||| |||||||||| |||
Db        660 AGAAGATGCAGCCTCCAGCCACCTGCAGTCAACCCATGCACCTTGGAGGCTCATGGGTTC 719

Qy        765 TGACTTCTGACC-TTCTCCTCTTCGCTCCCCC 795
              |||||||||||| ||||||||||||||||||
Db        720 TGACTTCTGACCTTTCTCCTCTTCGCTCCCCC 751


RESULT 13
BG164929
LOCUS       BG164929               1053 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602343555F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4453813 5',
            mRNA sequence.
ACCESSION   BG164929
VERSION     BG164929.1  GI:12671563
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1053)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10244  row: m  column: 14
            High quality sequence start: 3
            High quality sequence stop: 675.
FEATURES             Location/Qualifiers
     source          1. .1053
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4453813"
                     /tissue_type="hypernephroma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_89"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             77.9%;  Score 621;  DB 12;  Length 1053;
  Best Local Similarity   94.1%;  Pred. No. 3.9e-118;
  Matches  722;  Conservative    0;  Mismatches   35;  Indels   10;  Gaps    7;

Qy         20 CGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAG 79
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db          2 CGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGG 61

Qy         80 CCCCAG-CCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCA 138
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCCAGTCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCA 121

Qy        139 CGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTG 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTG 181
```

```
Qy        199 GCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCT 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCT 241

Qy        259 TGGCTGGGACCTATCTGAACTACCTGGGCCCCCC----TTTCAACGAGCCAGACTTCAAC 314
              |||||||||||||||||||||| || || |||||    |          | ||||||||||
Db        242 TGGCTGGGACCTATCTGAACTAGCTTGGGCCCCCCGTTTCAACTGAGCCTAGACTTCAAC 301

Qy        315 CCTCCCCGCCTGGGGGCAGAGACTCTG-CCCAGGGCCACTGTTGACTTGGAGGTGTGGCG 373
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        302 CCTCCCCGCCTGGGGGCAGAGACTCTGTCCCAGGGCCACTGTTGACTTGGAGGTGTGGCG 361

Qy        374 AAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTG 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 AAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTG 421

Qy        434 TTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCA 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCA 481

Qy        494 CTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGG 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGG 541

Qy        554 CTACCCACTG-CCCCAGCCGCTGCCTGGGACTGAACCCAC-TTGGACTCCTGGCCCTGCC 611
              |||||||||| |||||||||||||||||||||||||||| |||||||||||||||   |
Db        542 CTACCCACTGCCCCCAGCCGCTGCCTGGGACTGAACCCACTTTGGACTCCTGGCCCTTGC 601

Qy        612 CACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGG 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGG 661

Qy        672 CTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCT 731
              |||||||||||||||| |||||||| |||||  ||     |   ||||||||||||||| |
Db        662 CTGTGGCGCTCGGCC-AGGACTTCCACCGGTCCAGACGAAGATTGCAGCCTCCAGCAGTT 720

Qy        732 GCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||  ||| || ||||  ||||||| ||
Db        721 GCAGTCACCCTGGACCTGGGGGGCCAT-GCGTCTGGGTTCTGACTTT 766
```

RESULT 14
CF271930/c

```
LOCUS       CF271930                801 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AGENCOURT_15196924 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002178 5', mRNA sequence.
ACCESSION   CF271930
VERSION     CF271930.2  GI:38453381
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 801)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627842.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 08
            High quality sequence start: 6
            High quality sequence stop: 646.
```

```
FEATURES             Location/Qualifiers
     source          1. .801
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002178"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 75.9%; Score 604.8; DB 14; Length 801;
Best Local Similarity 99.2%; Pred. No. 7.9e-115;
Matches 606; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        185 GGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCA 244
              ||||||||||||||||||||||||||||||||||||| |||||| |||||||||||| |||||||
Db        702 GGACCCAGGGCCTGGCCCCTCCATCCAGAAAACNTATGACTTCACCCGCTACNTGGAGCA 643

Qy        245 CCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACGAGCC 304
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        642 CCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACNTGGGCCCCCCCTTTCAACGAGCC 583

Qy        305 AGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGA 364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 AGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGA 523

Qy        365 GGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCA 424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 GGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCA 463

Qy        425 CCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAG 484
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        462 CCTTCTGTGTTACTTGCGTGGCCTCAACCGTCGGGCTGCCACTGCTGAGCTGCGCCGCAG 403

Qy        485 CCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGC 544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 CCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGC 343

Qy        545 AGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGG 604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 AGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGG 283

Qy        605 CCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCA 664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 CCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCA 223

Qy        665 GACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCC 724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        222 GACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCC 163

Qy        725 AGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTC 784
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        162 AGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTC 103

Qy        785 TTCGCTCCCCC 795
```

```
                ||||||||||||
Db        102 TTCGCTCCCCC 92


RESULT 15
CF271924/c
LOCUS       CF271924                763 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AGENCOURT_15197020 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002181 5', mRNA sequence.
ACCESSION   CF271924
VERSION     CF271924.2  GI:38453378
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627836.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 11
            High quality sequence start: 7
            High quality sequence stop: 648.
FEATURES              Location/Qualifiers
     source           1. .763
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:7002181"
                      /tissue_type="mixed"
                      /lab_host="DH5A (T1 phage-resistant)"
                      /clone_lib="NIH_MGC_195"
                      /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                      loxP-HindIII; Clones from this library have been
                      PCR-amplified using gene-specific primers to contain the
                      complete open reading frame (based on known gene sequences
                      available from NCBI's RefSeq). Template for PCR is cDNA
                      derived from either pooled cytoplasmic polyA RNA from 30
                      cells lines or pooled total RNA from 10 different tissues
                      (from BD Biosciences/Clontech and Washington University).
                      PCR products are directionally cloned into the loxP sites
                      of the pDNR-Dual vector. Library constructed by Dr.
                      Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                      Expression Laboratory, Research Technology Program, SAIC
                      Frederick, NCI-Frederick, Frederick, MD 21702). For
                      information on which gene each clone represents, please
                      visit our anonymous ftp site at
                      ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                      a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match          75.5%;  Score 602;  DB 14;  Length 763;
  Best Local Similarity  98.9%;  Pred. No. 2.9e-114;
  Matches  617;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    1;

Qy        172 TCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCC 231
              ||||||||||||||||  ||  ||||||||||||||||||||||||||||||||||||||
Db        713 TCAATCGCACAGGGG--CCCAGGCCTGGCCCCTCCATCCAGAAAACCTATGACTTCACCC 656

Qy        232 GCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCC 291
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 GCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCC 596

Qy        292 CTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCA 351
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 CTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCA 536

Qy        352 CTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACG 411
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 CTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACG 476

Qy        412 AGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTG 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 AGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTG 416

Qy        472 AGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTG 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 AGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTG 356

Qy        532 CGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCA 591
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCA 296

Qy        592 CTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGC 651
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGC 236

Qy        652 TGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGA 711
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGA 176

Qy        712 AGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTC 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 AGATGCAGCCTCCAGCAGCTGCAGTCACCCCGCACCTGGGGGCTCATGGCTTCTGACTTC 116

Qy        772 TGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||
Db        115 TGACCTTCTCCTCTTCGCTCCCCC 92


Search completed: August 16, 2004, 04:01:32
Job time : 1823.54 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 12:15:49 ; Search time 272.88 Seconds
(without alignments)
12407.700 Million cell updates/sec

Title: US-09-931-704-1
Perfect score: 797
Sequence: 1 attaaagcttcgccggagcc.....tctcctcttcgctccccccc 797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	797	100.0	797	2	AAV22652	Aav22652 cDNA enco
	2	797	100.0	797	2	AAV47510	Aav47510 Human neu
	3	797	100.0	797	3	AAA39481	Aaa39481 Human NNT
	4	797	100.0	797	6	ABK11647	Abk11647 Human cDN
	5	793.4	99.5	881	4	AAH74484	Aah74484 Nucleotid
	6	793.4	99.5	1790	3	AAA88546	Aaa88546 Human int
	7	759.2	95.3	768	4	AAH99772	Aah99772 Human pro
	8	749.4	94.0	1710	2	AAX16161	Aax16161 Human car
	9	710	89.1	729	4	AAD04201	Aad04201 Human car
	10	694.8	87.2	1008	4	AAK51548	Aak51548 Human pol
	11	669.4	84.0	819	2	AAV22654	Aav22654 cDNA enco
	12	669.4	84.0	819	2	AAV47512	Aav47512 Mouse neu
	13	669.4	84.0	819	3	AAA39483	Aaa39483 Murine NN
	14	669.4	84.0	819	6	ABK11649	Abk11649 Mouse cDN
	15	668.4	83.9	968	4	ABA09140	Aba09140 Human car
	16	668.4	83.9	968	4	AAK52532	Aak52532 Human pol
	17	558.4	70.1	648	3	AAA88547	Aaa88547 Mouse int
	18	523.4	65.7	5087	3	AAA39482	Aaa39482 Human NNT
	19	523.4	65.7	5087	6	ABK11648	Abk11648 Human nov
	20	523.4	65.7	5088	2	AAV22653	Aav22653 Human gen
	21	523.4	65.7	5088	2	AAV47511	Aav47511 Human neu
c	22	493.4	61.9	495	4	AAI25564	Aai25564 Probe #15
c	23	493.4	61.9	495	4	ABA71773	Aba71773 Human foe
c	24	493.4	61.9	495	4	AAI52108	Aai52108 Probe #20
c	25	493.4	61.9	495	4	ABA37855	Aba37855 Probe #16
c	26	493.4	61.9	495	4	AAK46202	Aak46202 Human bon
c	27	493.4	61.9	495	4	AAK20142	Aak20142 Human bra
c	28	493.4	61.9	495	4	ABS45921	Abs45921 Human liv
c	29	493.4	61.9	495	6	ABS20513	Abs20513 Human gen
c	30	369	46.3	492	4	AAI16384	Aai16384 Probe #63
c	31	369	46.3	492	4	ABA59256	Aba59256 Human foe
c	32	369	46.3	492	4	AAI39047	Aai39047 Probe #77
c	33	369	46.3	492	4	ABA27996	Aba27996 Probe #64
c	34	369	46.3	492	4	AAK33253	Aak33253 Human bon
c	35	369	46.3	492	4	AAK07469	Aak07469 Human bra
c	36	369	46.3	492	4	ABS33003	Abs33003 Human liv
c	37	369	46.3	492	6	ABS08085	Abs08085 Human gen
	38	269.4	33.8	283	4	AAH23145	Aah23145 Osteoarth
	39	80.2	10.1	396	2	AAX16162	Aax16162 Human car
	40	73.6	9.2	1260	6	ABQ16903	Abq16903 Oligonucl
c	41	73.6	9.2	1260	6	ABQ16902	Abq16902 Oligonucl
	42	58.6	7.4	65	6	ABN56698	Abn56698 Mouse spl
	43	49	6.1	627	2	AAX60797	Aax60797 Human DNA
	44	46.8	5.9	1260	6	ABQ16904	Abq16904 Oligonucl
c	45	46.8	5.9	1260	6	ABQ16905	Abq16905 Oligonucl

ALIGNMENTS

RESULT 1
AAV22652
ID AAV22652 standard; cDNA; 797 BP.
XX
AC AAV22652;
XX
DT 13-JUL-1998 (first entry)
XX
DE cDNA encoding human neurotrophic factor NNT-1.
XX
KW Human; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW treatment; neurological disease; degeneration; Parkinson's disease;
KW amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 90. .767
FT /*tag= a
FT sig_peptide 90. .170
FT /*tag= b
FT mat_peptide 171. .764
FT /*tag= c
XX
PN US5741772-A.
XX
PD 21-APR-1998.
XX
PF 03-FEB-1997; 97US-00792019.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 1998-260526/23.
DR P-PSDB; AAW56141.
XX
PT Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT for stimulating growth of motor and sympathetic neurons.
XX
PS Disclosure; Fig 1; 41pp; English.
XX
CC The present sequence encodes a human neurotrophic factor, designated NNT-
CC 1, which is capable of stimulating growth of motor or sympathetic

```
CC   neurons. The NNT-1 protein is useful in the treatment of neurological
CC   diseases characterised by the degeneration and death of particular
CC   classes of neurons. These diseases specifically include Parkinson's
CC   disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke
CC   and various degenerative disorders affecting vision
XX
SQ   Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 797;  DB 2;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-176;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797


RESULT 2
AAV47510
ID   AAV47510 standard; cDNA; 797 BP.
XX
AC   AAV47510;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Human neurotrophic factor NNT-1 cDNA.
XX
KW   NNT-1; neurotrophic factor; human; antiinflammatory; adjuvant;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   peripheral neurophathy; dystrophy; neural retina degeneration;
KW   common variable immunodeficiency; CVID; selective IgA deficiency;
KW   hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             90. .767
FT                   /*tag=  a
FT   sig_peptide     90. .170
FT                   /*tag=  b
FT   mat_peptide     171. .764
FT                   /*tag=  c
XX
PN   WO9833922-A1.
XX
PD   06-AUG-1998.
XX
PF   02-FEB-1998;   98WO-US002363.
XX
PR   03-FEB-1997;   97US-00792019.
PR   30-JAN-1998;   98US-00016534.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M,  Elliot GS,  Senaldi G,  Sarmiento U;
XX
DR   WPI; 1998-437475/37.
DR   P-PSDB; AAW29715.
XX
PT   Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT   NNT-1 - useful for treatment of neurological and immunological diseases
PT   or inflammation, also as vaccine adjuvant.
XX
PS   Claim 3; Fig 1; 120pp; English.
XX
CC   This newly isolated human cDNA sequence (deposited at ATCC 98295) codes
CC   for a novel neurotrophic factor, designated NNT-1 (see AAW29715), that is
CC   a growth factor for neurons and for B or T cells. It was obtained from a
CC   T-cell lymphoma cDNA library by expressed sequence tag analysis on the
CC   basis of homology to CNTF. The isolated NNT-1 cDNA was used as probe to
CC   isolated NNT-1 genomic DNA (see AAV47511). Vectors containing the cDNA or
CC   genomic DNA and host cells are provided for use in the production of NNT-
CC   1 polypeptides. These are used to treat: (i) neurological or
CC   immunological diseases, specifically Alzheimer's, Parkinson's or
CC   Huntington's diseases, amyotrophic lateral sclerosis, Charcot-Marie-Tooth
CC   syndrome, peripheral neuropathy, dystrophy and degeneration of the neural
CC   retina, or conditions characterised by T or B cell defects, e.g. common
CC   variable immunodeficiency (CVID), selective IgA deficiency,
CC   hypogammaglobulinaemia and X-linked agammaglobulinaemia (claimed), but
CC   many others disclosed; and (ii) inflammation. NNT-1 is also able to boost
CC   immunoreactivity and antibody production following vaccination, and,
CC   since it inhibits tumour necrosis factor production, it may also be
CC   useful for treating sepsis. NNT-1 nucleic acid fragments are also used as
CC   hybridisation probes in diagnostic assays. In addition, cells that have
CC   been engineered to express NNT-1 can be implanted, or nucleic acids are
CC   delivered in gene therapy vectors
XX
SQ   Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 797;  DB 2;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-176;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 3
AAA39481

```
ID   AAA39481 standard; cDNA; 797 BP.
XX
AC   AAA39481;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Human NNT-1 cDNA.
XX
KW   NNT-1; human; neurotrophic factor; nootropic; neuroprotective; treatment;
KW   anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW   nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW   retinopathy; immune disorder; hematopoietic disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             90. .767
FT                   /*tag= a
FT                   /product= "NNT-1"
XX
PN   US6054294-A.
XX
PD   25-APR-2000.
XX
PF   12-DEC-1997;   97US-00988819.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 2000-338492/29.
DR   P-PSDB; AAY87813.
XX
PT   New nucleic acids encoding neurotrophic factors useful for stimulating
PT   growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS   Claim 1a; Fig 1; 42pp; English.
XX
CC   This invention describes a novel nucleic acid molecule (I) encoding a
CC   novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC   neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC   ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC   which are useful for treating patients in whom various cells of the
CC   central, autonomic, or peripheral nervous system have degenerated and/or
CC   have been damaged by congenital disease, trauma, mechanical damage,
CC   surgery, stroke, ischemia, infection, metabolic disease, nutritional
CC   deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC   treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC   sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC   neuropathy induced by diabetes or other metabolic disorders, and/or
CC   dystrophies or degeneration of the neural retina such as retinitis
CC   pigmentosa, drug-induced retinopathies, stationary forms of night
CC   blindness, progressive cone-rod degeneration, immune disorders and
CC   hematopoietic disorders. (I) is effective in treating neurological
CC   conditions and promotes neuron regeneration. Neural functions are
CC   effectively restored in patients suffering from various neurological
CC   disorders. This sequence encodes the human NNT-1 protein described in the
CC   method of the invention
XX
SQ   Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 797;  DB 3;  Length 797;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-176;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
```

```
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 4
ABK11647
ID ABK11647 standard; cDNA; 797 BP.
XX
AC ABK11647;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human cDNA encoding novel neurotrophic factor NNT1.
XX
KW Human; ss; gene; NNT1; neurotrophic factor; IgE-related disease;
KW Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW vascular restenosis; rheumatoid arthritis; psoriatic arthritis;
KW inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW inflammatory bowel disease; transplant rejection; reproductive disorder;
KW graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   90. .767
FT   /*tag= a
FT   /product= "NNT1"
XX
PN WO200215977-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-US025906.
XX
PR 18-AUG-2000; 2000US-0226436P.
PR 16-AUG-2001; 2001US-00931704.
XX
PA (AMGE-) AMGEN INC.
XX
PI Senaldi G;
XX
DR WPI; 2002-280867/32.
DR P-PSDB; AAU78176.
XX
PT Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT patient, preventing IgE-related disease and treating allergic diseases,
PT involves administering NNT-1 inhibitor to a patient.
XX
PS Claim 2; Fig 1; 63pp; English.
XX
CC The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC modulating IgE levels in a patient, preventing an IgE-related disease,
CC and treating allergic diseases, comprising administering a
CC therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC related disease or susceptibility to an IgE-related disease, by
CC determining the presence or amount of expression of an NNT1 polypeptide
CC encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC occurring variant, and diagnosing an IgE-related disease or
CC susceptibility of an IgE-related disease based on the presence or amount
CC of expression of the polypeptide and a pharmaceutical composition for use
CC in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC inhibitor is useful for preventing and treating IgE-related disease,
CC modulating IgE levels, and treating allergic diseases e.g. Type I
CC allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC asthma, immune diseases and disorders, diseases involving abnormal cell
CC proliferation including cancer, arteriosclerosis and vascular restenosis,
CC diseases and conditions relating to dysfunction of immune system
CC including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC transplant rejection, and graft versus host disease, and reproductive
CC diseases and disorders including infertility, miscarriage, preterm labour
CC and delivery, and endometriosis. The present sequence encodes human NNT1
XX
SQ Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

Query Match 100.0%; Score 797; DB 6; Length 797;
Best Local Similarity 100.0%; Pred. No. 3.6e-176;
Matches 797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
```

```
Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 5
AAH74484
ID AAH74484 standard; DNA; 881 BP.
XX
AC AAH74484;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human NNT-1 protein.
XX
KW NNT-1; CLF-1; sCNTFRalpha; nervous system; neuron; nervous system;
KW neuro-muscular function; tumour; immune system; haematopoietic system;
KW reproductive system; liver; skeletal muscle; neurodegenerative disease;
KW amyotrophic lateral sclerosis; Parkinson's disease; Huntingdon's disease;
KW muscular mass; paralysis; cancer; obesity; fertility; endometriosis;
KW blastocyst implantation; thrombosis; retinal disease;
KW retinal pigmentosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 174. .851
FT /*tag= a
FT /product= "NNT-1"
XX
PN WO200155172-A2.
XX
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-FR000253.
XX
PR 27-JAN-2000; 2000FR-00001035.
PR 12-OCT-2000; 2000FR-00013089.
XX
PA (FABR) FABRE MEDICAMENT SA PIERRE.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Elson G, Gauchat J, Plun-Favreau H, Chevalier S, Gascan H;
XX
DR WPI; 2001-488773/53.
DR P-PSDB; AAG63543.
XX
PT A complex comprising a NNT-1 protein and a CLF-1 and/or sCNTFRalpha
PT protein useful to treat neurodegenerative disease including Parkinson's
PT and Huntington's, obesity and cancer.
XX
PS Disclosure; Page 57-58; 67pp; French.
XX
CC The present sequence encodes a human NNT-1 protein. The specification
CC describes a complex comprising a NNT-1 protein and a CLF-1 and/or
CC sCNTFRalpha protein. The NNT-1/CLF-1 complex is used to modulate activity
CC of the sCNTFRalpha/gp130/LIFRbeta receptor complex, or to induce
CC phosphorylation of the tyrosine of gp130 and LIFRbeta, particularly where
CC cells expressing the receptor complex are in the central or peripheral
CC nervous system, in neurons implicated in neuro-muscular function or in
CC skeletal muscle. The complex or antibodies are also used to decrease the
CC survival, growth or proliferation of tumour cells or to facilitate the
CC proliferation and/or inhibit differentiation of cells stocks. The complex
CC is also used to modulate activity of the gp130/LIFRbeta receptor or cells
CC expressing that receptor, particularly those cells implicated in the
CC immune, haematopoietic, nervous or reproductive system, the liver or
CC skeletal muscle. Molecules of the invention may be used to prevent or
CC treat neurodegenerative diseases including amyotrophic lateral sclerosis,
CC Parkinson's and Huntington's disease, to repair or regenerate nervous or
CC muscular tissue or to maintain muscular mass in paralysis patients. They
CC may also be used to treat cancer, obesity and associated diseases, and to
CC improve fertility, particularly to avoid endometriosis and/or assist
CC blastocyst implantation, thrombosis, or retinal disease, particular
CC retinal pigmentosis
XX
SQ Sequence 881 BP; 158 A; 318 C; 246 G; 159 T; 0 U; 0 Other;

Query Match 99.5%; Score 793.4; DB 4; Length 881;
Best Local Similarity 99.9%; Pred. No. 2.6e-175;
Matches 794; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 144

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        145 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 204

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 264

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 324

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 384

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 444

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 504

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 564

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 624

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 684
```

```
Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 744

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 804

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        805 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 864

Qy        781 CCTCTTCGCTCCCCC 795
              |||||||||||||||
Db        865 CCTCTTCGCTCCCCC 879
```

RESULT 6
AAA88546
ID AAA88546 standard; DNA; 1790 BP.
XX
AC AAA88546;
XX
DT 22-JAN-2001 (first entry)
XX
DE Human interleukin-B60 (IL-B60) gene.
XX
KW Interleukin-B60; IL-B60; human; cytokine; chromosome 11;
KW cytokine-like factor-1; haematopoietic; inflammation; antiinflammatory;
KW autoimmune disease; therapy; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	162. .809
FT		/*tag= a
FT	sig_peptide	162. .212
FT		/*tag= b
FT	mat_peptide	213. .806
FT		/*tag= c

XX
PN WO200053631-A1.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US006182.
XX
PR 11-MAR-1999; 99US-00267901.
XX
PA (SCHE) SCHERING CORP.
XX
PI Oppmann B, Timans JC, Kastelein RA, Bazan JF;
XX
DR WPI; 2000-587426/55.
DR P-PSDB; AAB19686.
XX
PT Cytokine-like factor 1 (CLF-1) and interleukin (IL)-B60 complexes,
PT polypeptides, and nucleic acids, useful in research, diagnosis and for
PT treating inflammatory and autoimmune disorders.
XX
PS Claim 17; Page 15-16; 97pp; English.
XX
CC The present sequence is that of DNA encoding human interleukin-B60 (IL-
CC B60, see AAB19586), a novel, small soluble cytokine-like protein of 198
CC amino acids that exhibits structural motifs characteristic of a member of
CC the long-chain cytokines, and which shows homology to granulocyte colony
CC stimulating factor and interleukin-6. IL-60B may have either stimulatory
CC or inhibitory effects on haematopoietic cells, including e.g. lymphoid
CC cells, such as T-cells, B-cells, natural killer cells, macrophages,
CC dendritic cells, haematopoietic progenitors, etc. Methods are provided
CC for modulating the physiology or development of a cell or tissue culture
CC cells by contacting the cell with an agonist or antagonist of IL-B60 or
CC an agonist of antagonist of a complex of mature IL-B60 and its partner,
CC cytokine-like factor-1 (CLF-1, see AAB19588). The IL-B60/CLF-1 cytokine
CC serves as a key physiological factor in motor neuron development and
CC regeneration. IL-60B, its agonists and antagonists may be used to treat
CC inflammatory or autoimmune disorders and also for drug screening. The
CC IL60B gene maps to human chromosome 11
XX
SQ Sequence 1790 BP; 381 A; 560 C; 474 G; 375 T; 0 U; 0 Other;

Query Match 99.5%; Score 793.4; DB 3; Length 1790;
Best Local Similarity 99.9%; Pred. No. 3e-175;
Matches 794; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 102

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        103 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 162

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 222

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 282

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 342

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 402

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 462

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 522

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 582

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 642

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 702

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 762

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 822

Qy        781 CCTCTTCGCTCCCCC 795
              |||||||||||||||
Db        823 CCTCTTCGCTCCCCC 837
```

RESULT 7
AAH99772
ID AAH99772 standard; cDNA; 768 BP.
XX

AC AAH99772;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:607.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US035017.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457603/49.
DR P-PSDB; AAM25831.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS Claim 1; Page 638; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production, The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 768 BP; 137 A; 278 C; 217 G; 136 T; 0 U; 0 Other;

Query Match 95.3%; Score 759.2; DB 4; Length 768;
Best Local Similarity 99.6%; Pred. No. 2.3e-167;
Matches 761; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||| ||| ||||||||||||||||||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGTCCTTCCACTCCGCCAGCCTCCGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 124

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 184

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 244

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 304

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 364

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 424

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 484

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 544

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 604

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 664

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 724

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
              |||||||||||||||||||||||||||||||||||||||||||
Db        725 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 768
```

RESULT 8
AAX16161
ID AAX16161 standard; DNA; 1710 BP.
XX
AC AAX16161;
XX
DT 28-APR-1999 (first entry)
XX
DE Human cardiotrophin-like cytokine encoding DNA.
XX
KW Human; cardiotrophin-like cytokine; interleukin 6 cytokine family; CLC;
KW IL-6; diagnosis; detection; immune system-related disorder; cancer;
KW cardiac disorder; heart failure; hypertension; cancer;
KW autoimmune disorder; infection; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 46. .723

```
FT                   /*tag=  a
FT   sig_peptide     46. .129
FT                   /*tag=  b
FT   mat_peptide     130. .720
FT                   /*tag=  c
FT                   /product= "cardiotrophin-like cytokine"
XX
PN   WO9900415-A1.
XX
PD   07-JAN-1999.
XX
PF   29-JUN-1998;   98WO-US013129.
XX
PR   30-JUN-1997;   97US-0051311P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Shi Y,  Ruben SM;
XX
DR   WPI; 1999-095678/08.
DR   P-PSDB; AAW94466.
XX
PT   New isolated cardiotrophin-like cytokine nucleic acid - used to develop
PT   products for treating cardiac and immune system disorders, e.g. heart
PT   failure, hypertension, cancers, autoimmune disorders and infections.
XX
PS   Claim 2; Fig 1; 103pp; English.
XX
CC   The present invention relates to a novel cardiotrophin-like cytokine
CC   (CLC) protein which is a member of the interleukin 6 (IL-6) cytokine
CC   family. The present sequence encodes the human CLC protein. The present
CC   invention also describes screening methods for identifying agonists and
CC   antagonists of CLC activity, as well as methods for detecting cardiac and
CC   immune system-related disorders and therapeutic methods for treating
CC   cardiac and immune system-related disorders, e.g heart failure,
CC   hypertension, cancers, autoimmune disorders and infections
XX
SQ   Sequence 1710 BP; 370 A; 530 C; 448 G; 362 T; 0 U; 0 Other;

  Query Match             94.0%;  Score 749.4;  DB 2;  Length 1710;
  Best Local Similarity   99.9%;  Pred. No. 5.3e-165;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         45 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCA 104
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db          1 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCA 60

Qy        105 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 120

Qy        165 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 180

Qy        225 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 240

Qy        285 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 344
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 300

Qy        345 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 404
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 360

Qy        405 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 420

Qy        465 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 480

Qy        525 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 540

Qy        585 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 600

Qy        645 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 660

Qy        705 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 720

Qy        765 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||
Db        721 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 751


RESULT 9
AAD04201
ID   AAD04201 standard; cDNA; 729 BP.
XX
AC   AAD04201;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human cardiotrophin-like cytokine (CLC) cDNA.
XX
KW   Human; biologically active complex; haemopoietin receptor; NR6;
KW   cardiotrophin-like cytokine; CLC; therapy; prophylaxis; proliferation;
KW   differentiation; cell survival; neurotrophic activity; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .679
FT                   /*tag=  a
FT                   /product= "Human cardiotrophin-like cytokine (CLC)
FT                   protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
FT   sig_peptide     11. .91
FT                   /*tag=  b
FT   mat_peptide     92. .679
FT                   /*tag=  c
FT                   /product= "Human mature cardiotrophin-like cytokine (CLC)
FT                   protein"
XX
PN   WO200127157-A1.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-AU001216.
XX
PR   08-OCT-1999;   99AU-00003327.
PR   12-MAY-2000; 2000AU-00007489.
XX
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
PI   Nash A,  Jachno KM,  Fabri LJ,  Reid K,  Bartlett PF,  Hilton DJ;
PI   Nakata Y,  Hasegawa M;
XX
DR   WPI; 2001-281978/29.
DR   P-PSDB; AAE00828.
XX
PT   New biologically active complex comprising NR6 and cardiotrophin-like-
PT   cytokine, for facilitating proliferation, differentiation and/or survival
PT   of a cell.
```

```
XX
PS   Claim 31; Page 112-114; 123pp; English.
XX
CC   The present invention relates to a biologically active complex comprising
CC   a haemopoietin receptor, NR6 and cardiotrophin-like cytokine (CLC). The
CC   complex is useful in the manufacture of a medicament for the treatment
CC   and/or prophylaxis of a subject, as it is involved in facilitating
CC   proliferation, differentiation and/or survival of a cell. The complex or
CC   its components have neurotrophic activity. The present sequence is human
CC   cardiotrophin-like cytokine (CLC) cDNA
XX
SQ   Sequence 729 BP; 132 A; 261 C; 196 G; 140 T; 0 U; 0 Other;

  Query Match           89.1%;  Score 710;  DB 4;  Length 729;
  Best Local Similarity  100.0%;  Pred. No. 6.8e-156;
  Matches  710;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         86 CCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCT 145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCT 66

Qy        146 CTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTC 205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTC 126

Qy        206 CATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGG 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGG 186

Qy        266 GACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCT 325
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCT 246

Qy        326 GGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGA 385
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGA 306

Qy        386 CAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGG 445
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGG 366

Qy        446 CCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAG 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAG 426

Qy        506 CCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCC 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 CCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCC 486

Qy        566 CCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCT 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 CCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCT 546

Qy        626 CCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGC 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 CCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGC 606

Qy        686 CAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCA 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCA 666

Qy        746 CCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 CCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 716


RESULT 10
AAK51548
ID   AAK51548 standard; cDNA; 1008 BP.
XX
AC   AAK51548;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 93.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM78415.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 711-712; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 1008 BP; 183 A; 374 C; 247 G; 204 T; 0 U; 0 Other;

  Query Match           87.2%;  Score 694.8;  DB 4;  Length 1008;
  Best Local Similarity  97.0%;  Pred. No. 2.6e-152;
  Matches  708;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy         66 CCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTA 125
              ||||| | || |      | | |       ||| |   ||||||||||||||||||||||
Db        267 CCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTA 326

Qy        126 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 386

Qy        186 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 245
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        387 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCCCACCCGCTACCTGGAGCAC 446

Qy        246 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCA 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        447 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCA 506

Qy        306 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 566

Qy        366 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 626

Qy        426 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 686

Qy        486 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 746

Qy        546 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 806

Qy        606 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 866

Qy        666 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 926

Qy        726 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCT 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCT 986

Qy        786 TCGCTCCCCC 795
              ||||||||||
Db        987 TCGCTCCCCC 996
```

```
RESULT 11
AAV22654
ID   AAV22654 standard; cDNA; 819 BP.
XX
AC   AAV22654;
XX
DT   13-JUL-1998  (first entry)
XX
DE   cDNA encoding murine neurotrophic factor NNT-1.
XX
KW   Mouse; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW   treatment; neurological disease; degeneration; Parkinson's disease;
KW   amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             95. .771
FT                   /*tag=  a
FT   sig_peptide     95. .175
FT                   /*tag=  b
FT   mat_peptide     176. .769
FT                   /*tag=  c
FT
XX
PN   US5741772-A.
XX
PD   21-APR-1998.
XX
PF   03-FEB-1997;   97US-00792019.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 1998-260526/23.
DR   P-PSDB; AAW56142.
XX
PT   Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT   for stimulating growth of motor and sympathetic neurons.
XX
PS   Disclosure; Fig 4; 41pp; English.
XX
CC   The present sequence encodes a murine neurotrophic factor, designated NNT
CC   -1, which is capable of stimulating growth of motor or sympathetic
CC   neurons. The NNT-1 protein is useful in the treatment of neurological
CC   diseases characterised by the degeneration and death of particular
CC   classes of neurons. These diseases specifically include Parkinson's
CC   disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke
CC   and various degenerative disorders affecting vision
XX
SQ   Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 U; 0 Other;

  Query Match             84.0%;  Score 669.4;  DB 2;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 2e-146;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| |||| |||||||||| ||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| |||||||||||||||||  ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||  |||||||||||||||||||| ||||| |||||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||| |||||||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
```

```
                    |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783

RESULT 12
AAV47512
ID   AAV47512 standard; cDNA; 819 BP.
XX
AC   AAV47512;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Mouse neurotrophic factor NNT-1 cDNA.
XX
KW   NNT-1; neurotrophic factor; mouse; antiinflammatory; adjuvant;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   peripheral neurophathy; dystrophy; neural retina degeneration;
KW   common variable immunodeficiency; CVID; selective IgA deficiency;
KW   hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW   therapy; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             95. .772
FT                   /*tag=  a
FT   sig_peptide     95. .175
FT                   /*tag=  b
FT   mat_peptide     176. .769
FT                   /*tag=  c
XX
PN   WO9833922-A1.
XX
PD   06-AUG-1998.
XX
PF   02-FEB-1998;   98WO-US002363.
XX
PR   03-FEB-1997;   97US-00792019.
PR   30-JAN-1998;   98US-00016534.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M,  Elliot GS,  Senaldi G,  Sarmiento U;
XX
DR   WPI; 1998-437475/37.
DR   P-PSDB; AAW29716.
XX
PT   Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT   NNT-1 - useful for treatment of neurological and immunological diseases
PT   or inflammation, also as vaccine adjuvant.
XX
PS   Disclosure; Fig 4; 120pp; English.
XX
CC   This newly isolated mouse cDNA sequence codes for a novel neurotrophic
CC   factor, designated NNT-1 (see AAW29716), that is a growth factor for
CC   neurons and for B or T cells. Vectors and host cells are provided for use
CC   in the production of murine and human NNT-1 polypeptides. These are used
CC   to treat: (i) neurological or immunological diseases, specifically
CC   Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic lateral
CC   sclerosis, Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy
CC   and degeneration of the neural retina, or conditions characterised by T
CC   or B cell defects, e.g. common variable immunodeficiency (CVID),
CC   selective IgA deficiency, hypogammaglobulinaemia and X-linked
CC   agammaglobulinaemia (claimed), but many others disclosed; and (ii)
CC   inflammation. NNT-1 is also able to boost immunoreactivity and antibody
CC   production following vaccination, and, since it inhibits tumour necrosis
CC   factor production, it may also be useful for treating sepsis. NNT-1
CC   nucleic acid fragments are also used as hybridisation probes in
CC   diagnostic assays. In addition, cells that have been engineered to
CC   express NNT-1 can be implanted, or nucleic acids are delivered in gene
CC   therapy vectors
XX

SQ   Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 U; 0 Other;

  Query Match             84.0%;  Score 669.4;  DB 2;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 2e-146;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| ||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| |||||||||||||||| ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||  ||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| ||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| ||||||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              |||||||||||||| |||||| || |||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783

RESULT 13
AAA39483
ID   AAA39483 standard; cDNA; 819 BP.
XX
AC   AAA39483;
XX
DT   24-AUG-2000  (first entry)
XX
DE   Murine NNT-1 cDNA.
XX
KW   NNT-1; neurotrophic factor; nootropic; neuroprotective; treatment;
KW   anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW   nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome; murine;
```

```
KW   Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW   retinopathy; immune disorder; hematopoietic disorder; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             95. .772
FT                   /*tag=  a
FT                   /product= "NNT-1"
XX
PN   US6054294-A.
XX
PD   25-APR-2000.
XX
PF   12-DEC-1997;   97US-00988819.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 2000-338492/29.
DR   P-PSDB; AAY87814.
XX
PT   New nucleic acids encoding neurotrophic factors useful for stimulating
PT   growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS   Claim 2a; Fig 4; 42pp; English.
XX
CC   This invention describes a novel nucleic acid molecule (I) encoding a
CC   novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC   neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC   ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC   which are useful for treating patients in whom various cells of the
CC   central, autonomic, or peripheral nervous system have degenerated and/or
CC   have been damaged by congenital disease, trauma, mechanical damage,
CC   surgery, stroke, ischemia, infection, metabolic disease, nutritional
CC   deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC   treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC   sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC   neuropathy induced by diabetes or other metabolic disorders, and/or
CC   dystrophies or degeneration of the neural retina such as retinitis
CC   pigmentosa, drug-induced retinopathies, stationary forms of night
CC   blindness, progressive cone-rod degeneration, immune disorders and
CC   hematopoietic disorders. (I) is effective in treating neurological
CC   conditions and promotes neuron regeneration. Neural functions are
CC   effectively restored in patients suffering from various neurological
CC   disorders. This sequence encodes the murine NNT-1 protein described in
CC   the method of the invention
XX
SQ   Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 U; 0 Other;

  Query Match             84.0%;  Score 669.4;  DB 3;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 2e-146;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| ||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| ||||||||||||||||| ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||||  ||||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              ||  ||||| ||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              ||  |||||||||||||||| |||||||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 14
ABK11649
ID   ABK11649 standard; cDNA; 819 BP.
XX
AC   ABK11649;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Mouse cDNA encoding novel neurotrophic factor NNT1.
XX
KW   Mouse; ss; gene; NNT1; neurotrophic factor; IgE-related disease;
KW   Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW   pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW   vascular restenosis; rheumatoid arthritis; psoriatic arthritis;
KW   inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW   autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW   inflammatory bowel disease; transplant rejection; reproductive disorder;
KW   graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             95. .772
FT                   /*tag=  a
FT                   /product= "NNT1"
XX
PN   WO200215977-A2.
XX
PD   28-FEB-2002.
XX
PF   17-AUG-2001; 2001WO-US025906.
XX
PR   18-AUG-2000; 2000US-0226436P.
PR   16-AUG-2001; 2001US-00931704.
XX
PA   (AMGE-) AMGEN INC.
```

```
XX
PI   Senaldi G;
XX
DR   WPI; 2002-280867/32.
DR   P-PSDB; AAU78177.
XX
PT   Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT   patient, preventing IgE-related disease and treating allergic diseases,
PT   involves administering NNT-1 inhibitor to a patient.
XX
PS   Claim 2; Fig 4; 63pp; English.
XX
CC   The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC   modulating IgE levels in a patient, preventing an IgE-related disease,
CC   and treating allergic diseases, comprising administering a
CC   therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC   inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC   related disease or susceptibility to an IgE-related disease, by
CC   determining the presence or amount of expression of an NNT1 polypeptide
CC   encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC   occurring variant, and diagnosing an IgE-related disease or
CC   susceptibility of an IgE-related disease based on the presence or amount
CC   of expression of the polypeptide and a pharmaceutical composition for use
CC   in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC   inhibitor is useful for preventing and treating IgE-related disease,
CC   modulating IgE levels, and treating allergic diseases e.g. Type I
CC   allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC   asthma, immune diseases and disorders, diseases involving abnormal cell
CC   proliferation including cancer, arteriosclerosis and vascular restenosis,
CC   diseases and conditions relating to dysfunction of immune system
CC   including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC   arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC   disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC   transplant rejection, and graft versus host disease, and reproductive
CC   diseases and disorders including infertility, miscarriage, preterm labour
CC   and delivery, and endometriosis. The present sequence encodes mouse NNT1
XX
SQ   Sequence 819 BP; 156 A; 288 C; 218 G; 157 T; 0 U; 0 Other;

  Query Match             84.0%;  Score 669.4;  DB 6;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 2e-146;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| |||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| ||||||||||||||| ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||| ||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||

Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||||| |||||||||||||||||||||||||||||| || |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| ||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              |||||||||||||| ||||| || |||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 15
ABA09140
ID   ABA09140 standard; cDNA; 968 BP.
XX
AC   ABA09140;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human cardiotrophin-like cytokine homologue cDNA, SEQ ID NO:916.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US003800.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11896.
XX
PT   Human proteins and DNA encoding sequences useful for preventing, treating
PT   or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT   and cancer.
XX
PS   Claim 1; Page 793-794; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
```

CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention
XX
SQ Sequence 968 BP; 179 A; 353 C; 244 G; 192 T; 0 U; 0 Other;

Query Match 83.9%; Score 668.4; DB 4; Length 968;
Best Local Similarity 97.0%; Pred. No. 3.6e-146;
Matches 681; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy         66 CCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTA 125
              ||||| | || |    | | |     ||| |   ||||||||||||||||||||||||
Db        267 CCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTA 326

Qy        126 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 386

Qy        186 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 446

Qy        246 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACGAGCCA 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACGAGCCA 506

Qy        306 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 566

Qy        366 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 626

Qy        426 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 686

Qy        486 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 746
```

```
Qy        546 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 806

Qy        606 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 866

Qy        666 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 926

Qy        726 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGA 767
              |||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGA 968
```

Search completed: August 15, 2004, 16:49:01
Job time : 277.88 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 16:49:33 ; Search time 460.864 Seconds
(without alignments)
8485.292 Million cell updates/sec

Title: US-09-931-704-1
Perfect score: 797
Sequence: 1 attaaagcttcgccggagcc.....tctcctcttcgctccccccc 797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	797	100.0	797	9	US-09-931-704-1	Sequence 1, Appli
	2	759.2	95.3	768	13	US-10-296-115-607	Sequence 607, App
	3	749.4	94.0	1710	15	US-10-212-793-1	Sequence 1, Appli
	4	669.4	84.0	819	9	US-09-931-704-4	Sequence 4, Appli
	5	668.4	83.9	968	13	US-10-276-774-916	Sequence 916, App
	6	523.4	65.7	5087	9	US-09-931-704-3	Sequence 3, Appli
c	7	493.4	61.9	495	9	US-09-864-761-23175	Sequence 23175, A
c	8	369	46.3	492	9	US-09-864-761-6462	Sequence 6462, Ap
	9	269.4	33.8	283	9	US-09-765-231A-75	Sequence 75, Appl
	10	175.8	22.1	809	13	US-10-027-632-134702	Sequence 134702,
	11	175.8	22.1	809	16	US-10-027-632-134702	Sequence 134702,
	12	80.2	10.1	396	15	US-10-212-793-7	Sequence 7, Appli
	13	58.6	7.4	65	10	US-09-908-975-29446	Sequence 29446, A
	14	46.8	5.9	2320	13	US-09-925-298-266	Sequence 266, App
	15	46.8	5.9	2320	15	US-10-102-806-266	Sequence 266, App
	16	44.8	5.6	68750	14	US-10-014-717-1	Sequence 1, Appli
	17	44	5.5	536	17	US-10-338-110-119	Sequence 119, App
	18	43.6	5.5	3195	15	US-10-205-823-143	Sequence 143, App
	19	43.6	5.5	3195	15	US-10-285-976-54	Sequence 54, Appl
	20	43.6	5.5	3195	16	US-10-295-027-1127	Sequence 1127, Ap
	21	42	5.3	2085	17	US-10-311-623-18	Sequence 18, Appl
	22	42	5.3	2771	13	US-10-112-944-494	Sequence 494, App
	23	42	5.3	3206	13	US-10-112-944-6	Sequence 6, Appli
c	24	41.8	5.2	594	13	US-10-142-426-10	Sequence 10, Appl
c	25	41.8	5.2	594	15	US-10-123-155-10	Sequence 10, Appl
c	26	41.8	5.2	594	15	US-10-146-731-10	Sequence 10, Appl
c	27	41.8	5.2	594	15	US-10-140-472-10	Sequence 10, Appl
c	28	41.8	5.2	594	15	US-10-141-761-10	Sequence 10, Appl
c	29	41.8	5.2	594	15	US-10-142-885-10	Sequence 10, Appl
c	30	41.8	5.2	594	15	US-10-158-790-10	Sequence 10, Appl
c	31	41.8	5.2	594	16	US-10-137-871-10	Sequence 10, Appl
c	32	41.8	5.2	594	16	US-10-140-923-10	Sequence 10, Appl
c	33	41.8	5.2	594	16	US-10-141-756-10	Sequence 10, Appl
c	34	41.8	5.2	594	16	US-10-141-759-10	Sequence 10, Appl
c	35	41.8	5.2	594	16	US-10-140-805-10	Sequence 10, Appl
c	36	41.8	5.2	594	16	US-10-140-864-10	Sequence 10, Appl
	37	41.8	5.2	1581	16	US-10-369-493-43081	Sequence 43081, A
c	38	40.8	5.1	980	15	US-10-175-523-176	Sequence 176, App
c	39	40.6	5.1	536	17	US-10-338-110-119	Sequence 119, App
	40	40.4	5.1	1083	15	US-10-156-761-3132	Sequence 3132, Ap
c	41	40.4	5.1	9025608	15	US-10-156-761-1	Sequence 1, Appli
c	42	40.2	5.0	586	15	US-10-008-063-1	Sequence 1, Appli
c	43	40.2	5.0	586	15	US-10-152-363A-59	Sequence 59, Appl
c	44	40.2	5.0	690	13	US-10-087-192-983	Sequence 983, App
c	45	40.2	5.0	903	12	US-10-380-703-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-09-931-704-1
; Sequence 1, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
;  APPLICANT: Senaldi, Giorgio
;  TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
;  TITLE OF INVENTION:  Inhibitors
;  FILE REFERENCE: A-695
;  CURRENT APPLICATION NUMBER: US/09/931,704
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: US 60/226,436
;  PRIOR FILING DATE: 2000-08-18
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (90)..(764)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (171)..()
;   OTHER INFORMATION: Met at -27
US-09-931-704-1

  Query Match           100.0%;  Score 797;  DB 9;  Length 797;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-208;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
```

```
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 2
US-10-296-115-607
; Sequence 607, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 607
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-607

Query Match 95.3%; Score 759.2; DB 13; Length 768;
Best Local Similarity 99.6%; Pred. No. 1.2e-197;
Matches 761; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGTCCTTCCACTCCGCCAGCCTCCGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 124

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 184

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 244

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 304

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 364

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 424

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 484

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 544

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 604

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 664

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 724

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
              |||||||||||||||||||||||||||||||||||||||||||
Db        725 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 768
```

RESULT 3
US-10-212-793-1
; Sequence 1, Application US/10212793
; Publication No. US20030087395A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Cardiotrophin-Like Cytokine
; FILE REFERENCE: PF385D1C1
; CURRENT APPLICATION NUMBER: US/10/212,793
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 09/438,299
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/106,182
; PRIOR FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: US 60/051,311
; PRIOR FILING DATE: 1997-06-30

```
;  NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 1
;   LENGTH: 1710
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   LOCATION: CDS
;   OTHER INFORMATION: (46)..(720)
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: (46)..(126)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (127)..(720)
US-10-212-793-1

  Query Match             94.0%;  Score 749.4;  DB 15;  Length 1710;
  Best Local Similarity   99.9%;  Pred. No. 5.8e-195;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          45 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCA 104
               ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db           1 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCA 60

Qy         105 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 120

Qy         165 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 180

Qy         225 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 240

Qy         285 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 300

Qy         345 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 360

Qy         405 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 420

Qy         465 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 480

Qy         525 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 540

Qy         585 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 644
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 600

Qy         645 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 704
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 660

Qy         705 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 720

Qy         765 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
               |||||||||||||||||||||||||||||||
Db         721 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 751
```

```
RESULT 4
US-09-931-704-4
; Sequence 4, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
;  APPLICANT: Senaldi, Giorgio
;  TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using NF
;  TITLE OF INVENTION:  Inhibitors
;  FILE REFERENCE: A-695
;  CURRENT APPLICATION NUMBER: US/09/931,704
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: US 60/226,436
;  PRIOR FILING DATE: 2000-08-18
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 819
;   TYPE: DNA
;   ORGANISM: Murine
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (95)..(769)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (176)..()
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptid
;   LOCATION: (176)..(769)
;   OTHER INFORMATION:
;   NAME/KEY: sig_peptide
;   LOCATION: (95)..(175)
;   OTHER INFORMATION:
US-09-931-704-4

  Query Match             84.0%;  Score 669.4;  DB 9;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 4e-173;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy           1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
               |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db           5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy          61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
               |||| |||||||||||| | |||||||||||||||||||||||||||||||||||||||
Db          65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy         120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
               |||||||| ||||| |||||||||| |||||||||||||||||||||||| ||||||
Db         125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy         180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
               ||||| || ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db         185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy         240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
               ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db         245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy         300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
               ||||| |||||||| |||||  || |||||||||| |||||||||||||||  ||  ||
Db         305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy         360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
               ||||| |||||||||||||||||||  |||||||||||||||||| ||||| |||
Db         365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy         420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
               || ||||| |||||||||||||||||||||||||||||||||||| ||||| || ||
Db         425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy         480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
               || |||||||||||||||| |||||||||||||||||||||||||||| || |||
Db         485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544
```

```
Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | ||  |||||||||||||||||||||| |||||  |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              |||||||||||||| |||||  || |||||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||| ||| ||| |||||  |||||||  |||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783
```

RESULT 5
US-10-276-774-916
; Sequence 916, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 916
; LENGTH: 968
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-916

Query Match 83.9%; Score 668.4; DB 13; Length 968;
Best Local Similarity 97.0%; Pred. No. 7.6e-173;
Matches 681; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy         66 CCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTA 125
              |||||  | ||  |      |  | |      ||| |   ||||||||||||||||||||
Db        267 CCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTA 326

Qy        126 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 386

Qy        186 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 446

Qy        246 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCA 305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCA 506

Qy        306 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAG 566

Qy        366 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCAC 626

Qy        426 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 CTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGC 686
```

```
Qy        486 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCA 746

Qy        546 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 GCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGC 806

Qy        606 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 CCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAG 866

Qy        666 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        867 ACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCA 926

Qy        726 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGA 767
              |||||||||||||||||||||||||||||||||||||||||||
Db        927 GCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGA 968
```

RESULT 6
US-09-931-704-3
; Sequence 3, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
; APPLICANT: Senaldi, Giorgio
; TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using NF
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: A-695
; CURRENT APPLICATION NUMBER: US/09/931,704
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/226,436
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (137)..(138)
; OTHER INFORMATION: product = "INTERVENING UNSEQUENCED REGION OF >1KB"
US-09-931-704-3

Query Match 65.7%; Score 523.4; DB 9; Length 5087;
Best Local Similarity 99.8%; Pred. No. 3.7e-133;
Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        271 ATCTGAACTACCTGGGCCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330
              | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3363 AGCTGAACTACCTGGGCCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422

Qy        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482

Qy        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542

Qy        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602

Qy        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662

Qy        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
```

```
Qy        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782

Qy        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842

Qy        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||
Db       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
```

RESULT 7
US-09-864-761-23175/c
; Sequence 23175, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23175
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005849.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
; OTHER INFORMATION: SWISSPROT HIT: Q63086, EVALUE 8.00e-03
; OTHER INFORMATION: NT HIT: gi11439486, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AI752561.1, EVALUE 0.00e+00
US-09-864-761-23175

Query Match 61.9%; Score 493.4; DB 9; Length 495;
Best Local Similarity 99.8%; Pred. No. 4.9e-125;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        273 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 436

Qy        333 GAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 AAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 376

Qy        393 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 316

Qy        453 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 256

Qy        513 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 196

Qy        573 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 136

Qy        633 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 76

Qy        693 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 16

Qy        753 GCTCATGGCTTCTGA 767
              |||||||||||||||
Db         15 GCTCATGGCTTCTGA 1
```

RESULT 8
US-09-864-761-6462/c
; Sequence 6462, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03

```
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6462
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005849.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
US-09-864-761-6462

  Query Match             46.3%;  Score 369;  DB 9;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-91;
  Matches  369;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        427 TTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCC 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCC 433

Qy        487 TGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAG 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 TGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAG 373

Qy        547 CTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCC 313

Qy        607 CTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGA 253

Qy        667 CCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAG 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAG 193

Qy        727 CAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTT 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTT 133

Qy        787 CGCTCCCCC 795
              |||||||||
Db        132 CGCTCCCCC 124


RESULT 9
US-09-765-231A-75
; Sequence 75, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
;  APPLICANT: Searle/Monsanto
;  APPLICANT:  Phippard, Deborah
;  APPLICANT:  Vasanthakamur, Geetha
;  APPLICANT:  Dotson, Stanton
;  APPLICANT:  Ma, Xiao-Jun
;  TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
;  TITLE OF INVENTION:  vectors, and cells
;  FILE REFERENCE: SO-3221 PR
;  CURRENT APPLICATION NUMBER: US/09/765,231A
;  CURRENT FILING DATE:  2001-01-18
;  NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 75
;   LENGTH: 283
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-765-231A-75

  Query Match             33.8%;  Score 269.4;  DB 9;  Length 283;
  Best Local Similarity   99.3%;  Pred. No. 7.8e-64;
  Matches  281;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         37 CTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACC 96
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 CTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACC 60

Qy         97 TCCGAGCAGGGGACT-CGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTC 155
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCGAGCAGGGGACTGCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTC 120

Qy        156 CCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAA 215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAA 180

Qy        216 ACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTG 275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTG 240

Qy        276 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTC 318
              |||||||||||||||||||||||||||||||||||||||||||
Db        241 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTC 283


RESULT 10
US-10-027-632-134702
; Sequence 134702, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
```

```
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134702
;   LENGTH: 809
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134702

  Query Match            22.1%;  Score 175.8;  DB 13;  Length 809;
  Best Local Similarity  86.9%;  Pred. No. 3.2e-38;
  Matches  192;  Conservative    1;  Mismatches   28;  Indels    0;  Gaps    0;

Qy         66 CCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTA 125
              ||||| | || |      | | |       ||| |   ||||||||||||||||||||||
Db         17 CCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTA 76

Qy        126 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 136

Qy        186 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 245
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GACCCAGGGCCWGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 196

Qy        246 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGG 286
              |||||||||||||||||||||||||| |||    || | |
Db        197 CAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCG 237


RESULT 11
US-10-027-632-134702
; Sequence 134702, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134702
;   LENGTH: 809
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134702

  Query Match            22.1%;  Score 175.8;  DB 16;  Length 809;
  Best Local Similarity  86.9%;  Pred. No. 3.2e-38;
  Matches  192;  Conservative    1;  Mismatches   28;  Indels    0;  Gaps    0;

Qy         66 CCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTA 125
              ||||| | || |      | | |       ||| |   ||||||||||||||||||||||
Db         17 CCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTA 76

Qy        126 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 GCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGG 136

Qy        186 GACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 245
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 GACCCAGGGCCWGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCAC 196

Qy        246 CAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGG 286
              |||||||||||||||||||||||||| |||    || | |
Db        197 CAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCG 237


RESULT 12
US-10-212-793-7
; Sequence 7, Application US/10212793
; Publication No. US20030087395A1
; GENERAL INFORMATION:
;  APPLICANT: Shi et al.
;  TITLE OF INVENTION: Cardiotrophin-Like Cytokine
;  FILE REFERENCE: PF385D1C1
;  CURRENT APPLICATION NUMBER: US/10/212,793
;  CURRENT FILING DATE:  2002-08-07
;  PRIOR APPLICATION NUMBER: US 09/438,299
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: US 09/106,182
;  PRIOR FILING DATE: 1998-06-29
;  PRIOR APPLICATION NUMBER: US 60/051,311
;  PRIOR FILING DATE: 1997-06-30
;  NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 7
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (199)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (293)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (306)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (360)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (371)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (377)
;   OTHER INFORMATION: n equals any nucleotide
;   FEATURE:
;   NAME/KEY: Site
;   LOCATION: (383)
;   OTHER INFORMATION: n equals any nucleotide
US-10-212-793-7

  Query Match            10.1%;  Score 80.2;  DB 15;  Length 396;
  Best Local Similarity  96.5%;  Pred. No. 4.1e-12;
  Matches   82;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        711 AAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTT 770
```

```
Db          4 ACGAGGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTAACTT 63

Qy        771 CTGACCTTCTCCTCTTCGCTCCCCC 795

Db         64 CTGACCTTCTCCTCTTCGCTCCCCC 88

RESULT 13
US-09-908-975-29446
; Sequence 29446, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
;  APPLICANT: SHOSHAN, Avi
;  APPLICANT:  WASSERMAN, Alon
;  APPLICANT:  MINTZ, Eli
;  APPLICANT:  MINTZ, Liat
;  APPLICANT:  FAIGLER, Simchon
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29446
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-908-975-29446

  Query Match             7.4%;  Score 58.6;  DB 10;  Length 65;
  Best Local Similarity   93.8%;  Pred. No. 2.9e-06;
  Matches   61;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        651 CTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAG 710

Db          1 CTGAAGGAGCTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAG 60

Qy        711 AAGAT 715

Db         61 AAGAT 65

RESULT 14
US-09-925-298-266
; Sequence 266, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103
;  CURRENT APPLICATION NUMBER: US/09/925,298
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 266
;   LENGTH: 2320
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-298-266

  Query Match             5.9%;  Score 46.8;  DB 13;  Length 2320;
  Best Local Similarity   46.2%;  Pred. No. 0.0062;
  Matches  156;  Conservative    0;  Mismatches  182;  Indels    0;  Gaps    0;

Qy         11 CGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGG 70

Db         64 CGCAGAGTCTGCGGACCCGGCGCCGAGGCGGCCACCCGAGACGCGGCGCGCACGCTCCGG 123

Qy         71 CCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTG 130

Db        124 CCTGCGCCCGGCCCGGCCATGGCGGCCCCCCGCCCGTCTCCCGCGATCTCCGTTTCGGTC 183

Qy        131 CCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCC 190

Db        184 TCGGCTCCGGCTTTTTACGCCCCGCAGAAGAAGTTCGGCCCTGTGGTGGCCCCAAAGCCC 243

Qy        191 AGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACT 250

Db        244 AAAGTGAATCCCTTCCGGCCCGGGGACAGCGAGCCTCCCCCGGCACCCGGGGCCCAGCGC 303

Qy        251 CCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTT 310

Db        304 GCACAGATGGGCCGGGTGGGCGAGATTCCCCCGCCGCCCCCGGAAGACTTTCCCCTGCCT 363

Qy        311 CAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGG 348

Db        364 CCACCTCCCCTTGCTGGGGATGGCGACGATGCAGAGGG 401

RESULT 15
US-10-102-806-266
; Sequence 266, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 266
;   LENGTH: 2320
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-102-806-266

  Query Match             5.9%;  Score 46.8;  DB 15;  Length 2320;
  Best Local Similarity   46.2%;  Pred. No. 0.0062;
  Matches  156;  Conservative    0;  Mismatches  182;  Indels    0;  Gaps    0;

Qy         11 CGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGG 70

Db         64 CGCAGAGTCTGCGGACCCGGCGCCGAGGCGGCCACCCGAGACGCGGCGCGCACGCTCCGG 123

Qy         71 CCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGATGTTAGCGTG 130

Db        124 CCTGCGCCCGGCCCGGCCATGGCGGCCCCCCGCCCGTCTCCCGCGATCTCCGTTTCGGTC 183

Qy        131 CCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCC 190

Db        184 TCGGCTCCGGCTTTTTACGCCCCGCAGAAGAAGTTCGGCCCTGTGGTGGCCCCAAAGCCC 243

Qy        191 AGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAACT 250

Db        244 AAAGTGAATCCCTTCCGGCCCGGGGACAGCGAGCCTCCCCCGGCACCCGGGGCCCAGCGC 303

Qy        251 CCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTT 310

Db        304 GCACAGATGGGCCGGGTGGGCGAGATTCCCCCGCCGCCCCCGGAAGACTTTCCCCTGCCT 363
```

```
Qy        311 CAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGG 348
              | |||  |||   ||||||   | |||  |||  ||||
Db        364 CCACCTCCCCTTGCTGGGGATGGCGACGATGCAGAGGG 401
```

Search completed: August 16, 2004, 10:40:04
Job time : 467.864 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 16:10:45 ; Search time 58.7376 Seconds
(without alignments)
7530.032 Million cell updates/sec

Title: US-09-931-704-1
Perfect score: 797
Sequence: 1 attaaagcttcgccggagcc.....tctcctcttcgctccccccc 797

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	797	100.0	797	1	US-08-792-019B-1	Sequence 1, Appli
	2	797	100.0	797	3	US-08-988-819-1	Sequence 1, Appli
	3	797	100.0	797	3	US-09-016-534-1	Sequence 1, Appli
	4	749.4	94.0	1710	3	US-09-106-182-1	Sequence 1, Appli
	5	669.4	84.0	819	1	US-08-792-019B-4	Sequence 4, Appli
	6	669.4	84.0	819	3	US-08-988-819-4	Sequence 4, Appli
	7	669.4	84.0	819	3	US-09-016-534-4	Sequence 4, Appli
	8	523.4	65.7	5087	1	US-08-792-019B-3	Sequence 3, Appli
	9	523.4	65.7	5087	3	US-08-988-819-3	Sequence 3, Appli
	10	523.4	65.7	5087	3	US-09-016-534-3	Sequence 3, Appli
	11	80.2	10.1	396	3	US-09-106-182-7	Sequence 7, Appli
	12	44.8	5.6	68750	3	US-09-335-409-1	Sequence 1, Appli
	13	44.8	5.6	68750	4	US-09-568-102-1	Sequence 1, Appli
	14	44.8	5.6	68750	4	US-09-567-969-1	Sequence 1, Appli
	15	44.8	5.6	68750	4	US-09-568-480-1	Sequence 1, Appli
	16	44.8	5.6	68750	4	US-09-568-486-1	Sequence 1, Appli
	17	44.8	5.6	68750	4	US-09-568-472-1	Sequence 1, Appli
	18	44.8	5.6	68750	4	US-09-567-899-1	Sequence 1, Appli
c	19	42.2	5.3	1590	4	US-09-252-991A-10070	Sequence 10070, A
	20	42.2	5.3	1878	4	US-09-252-991A-9765	Sequence 9765, Ap
c	21	39.6	5.0	430	4	US-09-621-976-16656	Sequence 16656, A
	22	39	4.9	364	4	US-09-621-976-17202	Sequence 17202, A
c	23	38.6	4.8	364	4	US-09-621-976-17202	Sequence 17202, A
	24	38.6	4.8	6858	4	US-09-252-991A-1219	Sequence 1219, Ap
	25	38.4	4.8	71989	4	US-09-443-501A-2	Sequence 2, Appli
c	26	38	4.8	223	1	US-08-383-761-1	Sequence 1, Appli
c	27	38	4.8	223	1	US-08-824-277-1	Sequence 1, Appli
c	28	37.8	4.7	3129	3	US-09-387-695-1	Sequence 1, Appli
c	29	37.8	4.7	49272	1	US-08-614-770A-1	Sequence 1, Appli
	30	37.6	4.7	1071	4	US-09-252-991A-14206	Sequence 14206, A
c	31	37.6	4.7	2388	4	US-09-252-991A-14064	Sequence 14064, A
	32	37.4	4.7	900	4	US-09-252-991A-9282	Sequence 9282, Ap
	33	37.4	4.7	3729	4	US-09-252-991A-9272	Sequence 9272, Ap
c	34	37.2	4.7	1242	4	US-09-252-991A-9668	Sequence 9668, Ap
c	35	37.2	4.7	1521	4	US-09-252-991A-9746	Sequence 9746, Ap
	36	37.2	4.7	2511	4	US-09-252-991A-9494	Sequence 9494, Ap
	37	37.2	4.7	2583	4	US-09-252-991A-9541	Sequence 9541, Ap
c	38	37.2	4.7	3984	4	US-09-016-434-1199	Sequence 1199, Ap
c	39	37.2	4.7	4559	4	US-09-919-172-61	Sequence 61, Appl
c	40	37	4.6	468	4	US-09-252-991A-3528	Sequence 3528, Ap
	41	37	4.6	1449	4	US-09-252-991A-3558	Sequence 3558, Ap
	42	37	4.6	1698	4	US-09-252-991A-3551	Sequence 3551, Ap
c	43	37	4.6	2196	4	US-09-252-991A-3536	Sequence 3536, Ap
c	44	36.8	4.6	336	4	US-09-252-991A-12979	Sequence 12979, A
	45	36.8	4.6	1059	4	US-09-252-991A-12662	Sequence 12662, A

ALIGNMENTS

```
RESULT 1
US-08-792-019B-1
; Sequence 1, Application US/08792019B
; Patent No. 5741772
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  1840 DEHAVILLAND DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/792,019B
;      FILING DATE:  03-FEB-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  90..764
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  171..764
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  90..170
US-08-792-019B-1

  Query Match           100.0%;  Score 797;  DB 1;  Length 797;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-189;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

```
RESULT 2
US-08-988-819-1
; Sequence 1, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  90..764
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  171..764
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  90..170
US-08-988-819-1

  Query Match             100.0%;  Score 797;  DB 3;  Length 797;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-189;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```


```
Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
```

```
Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 3
US-09-016-534-1
; Sequence 1, Application US/09016534
; Patent No. 6143874
; GENERAL INFORMATION:
; APPLICANT: CHANG, MING-SHI
; APPLICANT: ELLIOTT, GARY S.
; APPLICANT: SARMIENTO, ULLA
; APPLICANT: SENALDI, GIORGIO
; TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: ONE AMGEN CENTER
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,534
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/792,019
; FILING DATE: 03-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-442B
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 90..764
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 171..764
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 90..170
US-09-016-534-1

```
  Query Match           100.0%;  Score 797;  DB 3;  Length 797;
  Best Local Similarity 100.0%;  Pred. No. 3.6e-189;
  Matches  797;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60

Qy         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGGA 120

Qy        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCA 180

Qy        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGG 240

Qy        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAACG 300

Qy        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGACT 360

Qy        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTACA 420

Qy        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCC 480

Qy        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCA 540

Qy        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTC 600

Qy        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGC 660

Qy        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGC 720

Qy        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCT 780

Qy        781 CCTCTTCGCTCCCCCCC 797
              |||||||||||||||||
Db        781 CCTCTTCGCTCCCCCCC 797
```

RESULT 4
US-09-106-182-1
; Sequence 1, Application US/09106182
; Patent No. 6046035
; GENERAL INFORMATION:
; APPLICANT: Shi, Yanggu
; APPLICANT: Ruben, Steve
; TITLE OF INVENTION: Cardiotrophin-Like Cytokine
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc
; STREET: 9410 Key West Ave

```
;       CITY:  Rockville
;       STATE:  MD
;       COUNTRY:  US
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/106,182
;       FILING DATE:  Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 60/051,053
;       FILING DATE:  30-JUN-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brookes, A. Anders
;       REGISTRATION NUMBER:  36,373
;       REFERENCE/DOCKET NUMBER:  PF385
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-309-8504
;       TELEFAX:  301-309-8439
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1710 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  46..720
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  46..126
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  127..720
US-09-106-182-1

  Query Match             94.0%;  Score 749.4;  DB 3;  Length 1710;
  Best Local Similarity   99.9%;  Pred. No. 3e-177;
  Matches  750;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          45 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCATGGACCTCCGAGCA 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCGGCCCCAGCCCCATGGACCTCCGAGCA 60

Qy         105 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTG 120

Qy         165 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGAC 180

Qy         225 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 284
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTG 240

Qy         285 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 344
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCC 300

Qy         345 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 404
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAG 360

Qy         405 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 464
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 AACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCC 420

Qy         465 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 524
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 ACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGC 480

Qy         525 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 584
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACT 540

Qy         585 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 644
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTC 600

Qy         645 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 704
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTC 660

Qy         705 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 764
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 AAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTC 720

Qy         765 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
               |||||||||||||||||||||||||||||||
Db         721 TGACTTCTGACCTTCTCCTCTTCGCTCCCCC 751


RESULT 5
US-08-792-019B-4
; Sequence 4, Application US/08792019B
; Patent No. 5741772
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  1840 DEHAVILLAND DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/792,019B
;      FILING DATE:  03-FEB-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  819 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  95..769
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  176..769
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  95..175
US-08-792-019B-4

  Query Match             84.0%;  Score 669.4;  DB 1;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.9e-157;
```

```
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| ||||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| ||||||||||||||| ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| ||||||||||||||||||||| ||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||||||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| |||||||||||||||||||||||||||||||| ||| |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| |||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||| ||||| || ||||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 6
US-08-988-819-4
; Sequence 4, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  819 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  95..769
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  176..769
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  95..175
US-08-988-819-4

  Query Match             84.0%;  Score 669.4;  DB 3;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.9e-157;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| |||||||||| ||||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| ||||||||||||||| ||  ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| ||||||||||||||||||||| ||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||||||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| |||||||||||||||||||||||||||||||| ||| |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| |||||||| ||  | ||| |
```

```
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 7
US-09-016-534-4
; Sequence 4, Application US/09016534
; Patent No. 6143874
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    APPLICANT:  ELLIOTT, GARY S.
;    APPLICANT:  SARMIENTO, ULLA
;    APPLICANT:  SENALDI, GIORGIO
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,534
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442B
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  819 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  95..769
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  176..769
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  95..175
US-09-016-534-4

  Query Match             84.0%;  Score 669.4;  DB 3;  Length 819;
  Best Local Similarity   92.0%;  Pred. No. 1.9e-157;
  Matches  717;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy          1 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCCGGGAGAGGAG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||

Db          5 ATTAAAGCTTCGCCGGAGCCGCGGCTCGCCCTCCCACTCCGCCAGCCTCTGGGAGAGGAG 64

Qy         61 CCGCACCCGGCCGGCCCAG-CCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 119
              |||| |||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db         65 CCGCGCCCGGCCGGCCCGGCCCCCAGCCCCATGGACCTCCGAGCAGGGGACTCGTGGGGG 124

Qy        120 ATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGC 179
              |||||||| ||||| ||||||||||| ||||||||||||||||||||||||||| ||||||
Db        125 ATGTTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGC 184

Qy        180 ACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 239
              ||||| || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        185 ACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTG 244

Qy        240 GAGCACCAACTCCGCAGCTTGGCTGGGACCTATCTGAACTACCTGGGCCCCCCTTTCAAC 299
              ||||| |||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db        245 GAGCATCAACTCCGCAGCTTAGCTGGGACCTACCTGAACTACCTGGGGCCCCCTTTCAAC 304

Qy        300 GAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTCTGCCCAGGGCCACTGTTGAC 359
              ||||| |||||||| ||||| || |||||||||| |||||||||||||||||||| || ||
Db        305 GAGCCTGACTTCAATCCTCCTCGACTGGGGGCAGAAACTCTGCCCAGGGCCACGGTCAAC 364

Qy        360 TTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGACCCAGAACTACGAGGCCTAC 419
              ||||| |||||||||||||||||||| |||||||||||||||||||| ||||| |||
Db        365 TTGGAAGTGTGGCGAAGCCTCAATGACAGGCTGCGGCTGACCCAGAACTATGAGGCGTAC 424

Qy        420 AGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGC 479
              || ||||| |||||||||||||||||||||||||||||||||||||| ||||| || ||
Db        425 AGTCACCTCCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACAGCTGAACTCCGA 484

Qy        480 CGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTC 539
              || |||||||||||||||| ||||||||||||||||||||||||||||||||||| |||
Db        485 CGTAGCCTGGCCCACTTCTGTACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCAGGTGTC 544

Qy        540 ATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACT 599
              |||||  | || |||||||||||||||||||| ||||| ||||||| ||  | ||| |
Db        545 ATGGCGACGCTTGGCTACCCACTGCCCCAGCCTCTGCCAGGGACTGAGCCAGCCTGGGCC 604

Qy        600 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAG 659
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        605 CCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGATGACTTCTGGCTGCTGAAGGAG 664

Qy        660 CTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAG 719
              ||||||||||||||| ||||| || |||||||||||||||||||| ||||||||||||||||
Db        665 CTGCAGACCTGGCTATGGCGTTCAGCCAAGGACTTCAACCGGCTTAAGAAGAAGATGCAG 724

Qy        720 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTT 778
              |||||||||||| |||||||||||||| ||| ||| ||||| ||||||| ||||||| |
Db        725 CCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCACATGGTTTCTGACCTCTGACCCT 783


RESULT 8
US-08-792-019B-3
; Sequence 3, Application US/08792019B
; Patent No. 5741772
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  1840 DEHAVILLAND DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/792,019B
```

```
;       FILING DATE:  03-FEB-1997
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  COOK, ROBERT R.
;       REGISTRATION NUMBER:  31,602
;       REFERENCE/DOCKET NUMBER:  A-442
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5087 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  137..138
;       OTHER INFORMATION:  /product= "INTERVENING UNSEQUENCED
;       OTHER INFORMATION:  REGION OF >1KB"
US-08-792-019B-3

  Query Match           65.7%;  Score 523.4;  DB 1;  Length 5087;
  Best Local Similarity  99.8%;  Pred. No. 6.2e-121;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422

Qy        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482

Qy        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542

Qy        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602

Qy        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662

Qy        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722

Qy        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782

Qy        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842

Qy        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||
Db       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887


RESULT 9
US-08-988-819-3
; Sequence 3, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5087 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  137..138
;       OTHER INFORMATION:  /product= "INTERVENING UNSEQUENCED
;       OTHER INFORMATION:  REGION OF >1KB"
US-08-988-819-3

  Query Match           65.7%;  Score 523.4;  DB 3;  Length 5087;
  Best Local Similarity  99.8%;  Pred. No. 6.2e-121;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422

Qy        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482

Qy        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542

Qy        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602

Qy        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662

Qy        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722

Qy        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782

Qy        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842

Qy        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||
Db       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
```

```
RESULT 10
US-09-016-534-3
; Sequence 3, Application US/09016534
; Patent No. 6143874
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    APPLICANT:  ELLIOTT, GARY S.
;    APPLICANT:  SARMIENTO, ULLA
;    APPLICANT:  SENALDI, GIORGIO
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,534
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442B
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5087 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  137..138
;      OTHER INFORMATION:  /product= "INTERVENING UNSEQUENCED
;      OTHER INFORMATION:  REGION OF >1KB"
US-09-016-534-3

  Query Match             65.7%;  Score 523.4;  DB 3;  Length 5087;
  Best Local Similarity   99.8%;  Pred. No. 6.2e-121;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422

Qy        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482

Qy        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542

Qy        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602

Qy        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662

Qy        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722

Qy        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782

Qy        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842

Qy        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||||||||||||||||||||||
Db       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887


RESULT 11
US-09-106-182-7
; Sequence 7, Application US/09106182
; Patent No. 6046035
;  GENERAL INFORMATION:
;    APPLICANT:  Shi, Yanggu
;    APPLICANT:  Ruben, Steve
;    TITLE OF INVENTION:  Cardiotrophin-Like Cytokine
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/106,182
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/051,053
;      FILING DATE:  30-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF385
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  396 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-106-182-7

  Query Match             10.1%;  Score 80.2;  DB 3;  Length 396;
  Best Local Similarity   96.5%;  Pred. No. 4.7e-11;
  Matches   82;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        711 AAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTT 770
              | || |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db          4 ACGAGGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTAACTT 63

Qy        771 CTGACCTTCTCCTCTTCGCTCCCCC 795
              |||||||||||||||||||||||||
Db         64 CTGACCTTCTCCTCTTCGCTCCCCC 88
```

RESULT 12
US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1

Query Match 5.6%; Score 44.8; DB 3; Length 68750;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 112; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy      456 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 515
            ||||  ||  |  |  ||||  |  |     |  |  ||  |  ||||  |  |  |  ||  |||
Db    57992 CAGGTAGCGAGTCGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGC 58051

Qy      516 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 575
            |||||  ||  |  ||  |  |  |||||  |  |  |  |  |  |      ||  ||    |  ||
Db    58052 GTGCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTCGAACCTCGCCGCCCTGCGC 58111

Qy      576 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 635
            |  ||  |  |  ||  |  |  |||  ||  |  |  |||  ||  |  |||  |  |
Db    58112 GCCCGGCTTCATGCCGCCGTGCCCGCTGCGGCTATCTATGGTGCGCTCGCCGAGATGGGG 58171

Qy      636 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCG 679
              |  |  |  |  |  |  |  |  ||  ||  |||||||||
Db    58172 CTTCAATACGGCCCGGCGTTGCGGGGGCTCGCCGAGCTGTGGCG 58215
```

RESULT 13
US-09-568-102-1
; Sequence 1, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-102-1

Query Match 5.6%; Score 44.8; DB 4; Length 68750;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 112; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy      456 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 515
            ||||  ||  |  |  ||||  |  |     |  |  ||  |  ||||  |  |  |  ||  |||
Db    57992 CAGGTAGCGAGTCGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGC 58051

Qy      516 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 575
            |||||  ||  |  ||  |  |  |||||  |  |  |  |  |  |      ||  ||    |  ||
Db    58052 GTGCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTCGAACCTCGCCGCCCTGCGC 58111

Qy      576 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 635
            |  ||  |  |  ||  |  |  |||  ||  |  |  |||  ||  |  |||  |  |
Db    58112 GCCCGGCTTCATGCCGCCGTGCCCGCTGCGGCTATCTATGGTGCGCTCGCCGAGATGGGG 58171

Qy      636 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCG 679
              |  |  |  |  |  |  |  |  ||  ||  |||||||||
Db    58172 CTTCAATACGGCCCGGCGTTGCGGGGGCTCGCCGAGCTGTGGCG 58215
```

RESULT 14
US-09-567-969-1
; Sequence 1, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-567-969-1

Query Match 5.6%; Score 44.8; DB 4; Length 68750;
Best Local Similarity 50.0%; Pred. No. 0.12;
Matches 112; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy      456 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 515
            ||||  ||  |  |  ||||  |  |     |  |  ||  |  ||||  |  |  |  ||  |||
Db    57992 CAGGTAGCGAGTCGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGC 58051

Qy      516 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 575
            |||||  ||  |  ||  |  |  |||||  |  |  |  |  |  |      ||  ||    |  ||
Db    58052 GTGCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTCGAACCTCGCCGCCCTGCGC 58111

Qy      576 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 635
            |  ||  |  |  ||  |  |  |||  ||  |  |  |||  ||  |  |||  |  |
Db    58112 GCCCGGCTTCATGCCGCCGTGCCCGCTGCGGCTATCTATGGTGCGCTCGCCGAGATGGGG 58171

Qy      636 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCG 679
              |  |  |  |  |  |  |  |  ||  ||  |||||||||
Db    58172 CTTCAATACGGCCCGGCGTTGCGGGGGCTCGCCGAGCTGTGGCG 58215
```

RESULT 15
US-09-568-480-1
; Sequence 1, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross

```
;  APPLICANT:  Cyr, Devon
;  APPLICANT:  Goerlach, Joern
;  TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
;  FILE REFERENCE: 4-30582A
;  CURRENT APPLICATION NUMBER: US/09/568,480
;  CURRENT FILING DATE:  2000-05-10
;  PRIOR APPLICATION NUMBER: 09/335,409
;  PRIOR FILING DATE: 1999-06-17
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 68750
;   TYPE: DNA
;   ORGANISM: Sorangium cellulosum
US-09-568-480-1

  Query Match             5.6%;  Score 44.8;  DB 4;  Length 68750;
  Best Local Similarity   50.0%;  Pred. No. 0.12;
  Matches  112;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy        456 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 515
              ||||  || | |   |||| | |     | |  || | |||| | | |  |  ||  |||
Db      57992 CAGGTAGCGAGTCGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGC 58051

Qy        516 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 575
              ||||| || | || | | |||||  | |  |   | | |     ||  ||   | ||
Db      58052 GTGCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTCGAACCTCGCCGCCCTGCGC 58111

Qy        576 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 635
              |  ||  | |  || |   | |  |||   ||  | |  |||   ||  | ||| |  |
Db      58112 GCCCGGCTTCATGCCGCCGTGCCCGCTGCGGCTATCTATGGTGCGCTCGCCGAGATGGGG 58171

Qy        636 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCG 679
              | | | | | |  |  | |  ||  ||  |||||||||
Db      58172 CTTCAATACGGCCCGGCGTTGCGGGGGCTCGCCGAGCTGTGGCG 58215


Search completed: August 16, 2004, 09:35:11
Job time : 62.7376 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:43:28 ; Search time 13.5 Seconds
(without alignments)
1603.192 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	124.5	10.2	203	2	JC4645	cardiotrophin-1 -
2	118.5	9.7	203	2	I49153	cardiotrophin-1 -
3	92	7.5	332	2	G86182	hypothetical prote
4	91.5	7.5	201	2	G02312	cardiotrophin-1 -
5	91.5	7.5	1182	2	I48378	hairless protein -
6	91	7.4	195	2	JH0680	ciliary neurotroph
7	89.5	7.3	1313	2	T38943	probable guanine n
8	89	7.3	733	2	D83588	conserved hypothet
9	87	7.1	955	2	T00247	zinc finger protei
10	87	7.1	1561	2	T00248	zinc finger protei
11	86.5	7.1	640	2	T34916	transferase - Stre
12	85.5	7.0	560	2	C38604	poly(3-hydroxyalka
13	84.5	6.9	389	2	F72511	probable cystathio
14	84	6.9	200	2	AD3633	hypothetical prote
15	84	6.9	411	2	F75439	probable cell wall
16	84	6.9	559	2	G71327	probable apolipopr
17	83	6.8	452	2	A35602	tailless (tll) pro
18	82.5	6.7	542	2	A82965	hypothetical prote
19	82.5	6.7	723	2	B38749	3-phosphatidylinos
20	82	6.7	1179	2	T04584	TMV resistance pro
21	81	6.6	479	2	AG2726	DNA photolyase [im
22	81	6.6	479	2	C97508	blue-light photore
23	80.5	6.6	353	2	AB1823	hypothetical prote
24	80.5	6.6	512	2	S21171	activin receptor S
25	80	6.5	200	1	UNRTCF	ciliary neurotroph
26	80	6.5	460	2	D75493	cell division cycl
27	80	6.5	2261	1	A42548	genome polyprotein
28	79	6.4	812	2	T34180	hypothetical prote
29	79	6.4	1220	2	AD0125	exodeoxyribonuclea
30	79	6.4	1621	2	T30200	protein-tyrosine k
31	78.5	6.4	390	2	T35509	hypothetical prote
32	78.5	6.4	426	2	AI0287	conserved hypothet
33	78	6.4	254	2	T08755	yes-associated pro
34	78	6.4	272	2	T35231	hypothetical prote
35	77.5	6.3	741	2	A83271	hypothetical prote
36	77.5	6.3	7463	2	T36248	CDA peptide synthe
37	77	6.3	322	2	T23891	hypothetical prote
38	77	6.3	344	2	T15457	hypothetical prote
39	77	6.3	520	2	B75332	extracellular solu
40	77	6.3	586	1	WMBEDE	65K early nonstruc
41	77	6.3	587	1	WMBETE	65K early nonstruc
42	77	6.3	733	2	F82965	hypothetical prote
43	76.5	6.2	310	2	A82734	conserved hypothet
44	76.5	6.2	338	2	T45394	hypothetical prote
45	76.5	6.2	522	2	D87123	conserved hypothet

ALIGNMENTS

RESULT 1
JC4645
cardiotrophin-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Jun-2000
C;Accession: JC4645
R;Ishikawa, M.; Saito, Y.; Miyamoto, Y.; Kuwahara, K.; Ogawa, E.; Nakagawa, O.; Harada,
Biochem. Biophys. Res. Commun. 219, 377-381, 1996
A;Title: cDNA cloning of rat cardiotrophin-1 (CT-1): Augmented expression of CT-1 gene i
A;Reference number: JC4645; MUID:96193659; PMID:8604995
A;Accession: JC4645
A;Molecule type: mRNA
A;Residues: 1-203 <ISH>
A;Cross-references: DDBJ:D78591; NID:g1256926; PIDN:BAA11427.1; PID:g1256927
C;Genetics:
A;Gene: CT-1
C;Keywords: cardiac muscle; cytokine; heart

Query Match 10.2%; Score 124.5; DB 2; Length 203;
Best Local Similarity 28.1%; Pred. No. 0.00045;
Matches 50; Conservative 24; Mismatches 85; Indels 19; Gaps 5;

```
Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:  | ||  | |:||||   |      :   | |
Db         27 IRQTHNLARLLTKYADQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | | |  : |: |       | |||     :: |  ::  |:
Db         84 ---SERLRQDAAALSALPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGY----PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     |:|:|:  |   :|   |   |  |:    :      |: |:  |  :|
Db        141 AALGAAARGPVPEPV-ATSALFTSNSAAGVFSAKVLGLHVCGLYGEWVSRTEGDLGQL 197
```

RESULT 2
I49153
cardiotrophin-1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49153
R;Pennica, D.; King, K.L.; Shaw, K.J.; Luis, E.; Rullamas, J.; Luoh, S.; Darbonne, W.C.;
Proc. Natl. Acad. Sci. U.S.A. 92, 1142-1146, 1995
A;Title: Expression cloning of cardiotrophin 1, a cytokine that induces cardiac myocyte
A;Reference number: I49153; MUID:95166785; PMID:7862649
A;Accession: I49153
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-203 <RES>
A;Cross-references: EMBL:U18366; NID:g710331; PIDN:AAC52173.1; PID:g710332
C;Genetics:
A;Gene: ctf1

C;Keywords: cytokine; glycoprotein
F;164/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 118.5; DB 2; Length 203;
Best Local Similarity 28.8%; Pred. No. 0.0016;
Matches 51; Conservative 21; Mismatches 88; Indels 17; Gaps 5;

```
Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197
```

RESULT 3
G86182
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: G86182
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86182
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <STO>
A;Cross-references: GB:AE005172; NID:g7211973; PIDN:AAF40444.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 7.5%; Score 92; DB 2; Length 332;
Best Local Similarity 26.4%; Pred. No. 0.83;
Matches 58; Conservative 24; Mismatches 72; Indels 66; Gaps 13;

```
Qy         23 PAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNY----------------- 64
Db        113 PSVTAGNLSGYP-PRPSF--TYDPGPYEQRQMESLLQQFIRERNPQIRPLPRLGLGSPVG 169

Qy         65 LGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQA 124
Db        170 LGPIRASPQFLQPRVAPP--PTSILD----TSRNRKAR------SKDGALAVVRG--RKV 215

Qy        125 ATAELRRSL-----------AHFCTSLQGLLGSIAGVMAALGYPLPQPLP---GTEPTWT 170
Db        216 RITEGSSSLYSLGRSWLKNGAHV-----GIQPQRSGIMK----PLPKPLPVDLTTETSVP 266

Qy        171 PGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQ 210
Db        267 DDPDEESADEDKEDEEAVKQL------SEKDL--LKRHIE 298
```

RESULT 4
G02312
cardiotrophin-1 - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 17-Jul-1998
C;Accession: G02312
R;Wood, W.I.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01035
A;Accession: G02312
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-201 <WOO>
A;Cross-references: EMBL:U43030; NID:g1151149; PID:g1151150
C;Genetics:
A;Gene: GDB:CTF1; CT-1
A;Cross-references: GDB:567078
A;Map position: 1p22-1p22

Query Match 7.5%; Score 91.5; DB 2; Length 201;
Best Local Similarity 25.9%; Pred. No. 0.49;
Matches 45; Conservative 24; Mismatches 92; Indels 13; Gaps 5;

```
Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLND 99
Db         27 IRQTHSLAHLLTKYAEQLLQEYVQLQGDPFGLPSFSPPRLPVAGL-SAPAPSHAGLPVHE 85

Qy        100 KLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVMAAL 154
Db         86 RLRL--DAAALAALPPLLDAVCRRQAELNPRAPRLLRRLEDAARQARALGAAVEALLAAL 143

Qy        155 GYPLPQPLPGTEP---TWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
Db        144 G--AANRGPRAEPPAATASAASATGVFPAKVLGLRVCGLYREWLSRTEGDLGQL 195
```

RESULT 5
I48378
hairless protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48378
R;Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994
A;Title: Structure and expression of the hairless gene of mice.
A;Reference number: I48378; MUID:94329587; PMID:8052649
A;Accession: I48378
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1182 <RES>
A;Cross-references: EMBL:Z32675; NID:g531706; PIDN:CAA83587.1; PID:g531707

Query Match 7.5%; Score 91.5; DB 2; Length 1182;
Best Local Similarity 23.7%; Pred. No. 4.6;
Matches 70; Conservative 25; Mismatches 99; Indels 101; Gaps 17;

```
Qy          3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPG-------PGP----------------- 38
Db        813 LRAGS--GLRKGL-----SLPLSPVRTRLSPPGALLWLQEPRPKHGFHLFQEHWRQGQPV 865

Qy         39 ---SIQKTYDLTRYLEHQLRSLAGTY--LNYLGPP------------FNEPDFNPPRLG 80
Db        866 LVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPPQPTNLDSTAFWEGFSHPETRP---- 921

Qy         81 AETLPRATVDLEVWRSLNDK-LRLTQN-----------YEAYSHLLCYL-RGLNRQAAT 126
Db        922 --KLDEGSV-LLLHRTLGDKDASRVQNLASSLPLPEYCAHQGKLNLASYLPLGLTLHPLE 978

Qy        127 AEL-----------RRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAH 175
Db        979 PQLWAAYGVNSHRGHLGTKNLCVEVSDLISILVHAEAQL------------PPWY--RAQ 1024

Qy        176 SDFLQKMD--DFWLLKELQTWLWR--SAKDFNRLKKKMQP--PAAAVTLHLGAHG 224
Db       1025 KDFLSGLDGEGLWSPGSQTSTVWHVFRAQDAQRIRRFLQMVCPAGAGTLEPGAPG 1079
```

RESULT 6
JH0680
ciliary neurotrophic factor - chicken

```
N;Alternate names: growth-promoting activity protein
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C;Accession: JH0680; PQ0057
R;Leung, D.W.; Parent, A.S.; Cachianes, G.; Esch, F.; Coulombe, J.N.; Nikolics, K.; Ecke
Neuron 8, 1045-1053, 1992
A;Title: Cloning, expression during development, and evidence for release of a trophic f
A;Reference number: JH0680; MUID:92304573; PMID:1610564
A;Accession: JH0680
A;Molecule type: mRNA
A;Residues: 1-195 <LEU>
A;Cross-references: GB:M80827; NID:g211822; PIDN:AAA48784.1; PID:g211823
A;Experimental source: eye
R;Eckenstein, F.P.; Esch, F.; Holbert, T.; Blacher, R.W.; Nishi, R.
Neuron 4, 623-631, 1990
A;Title: Purification and characterization of a trophic factor for embryonic peripheral
A;Reference number: PQ0057; MUID:90211978; PMID:2322465
A;Accession: PQ0057
A;Molecule type: protein
A;Residues: 155-166,'X',168-175 <ECK>
A;Experimental source: sciatic nerves
C;Comment: This is a neurotrophic protein.
C;Superfamily: ciliary neurotrophic factor
C;Keywords: growth factor

  Query Match             7.4%;  Score 91;  DB 2;  Length 195;
  Best Local Similarity   27.3%;  Pred. No. 0.53;
  Matches   51;  Conservative   21;  Mismatches   83;  Indels   32;  Gaps    9;

Qy         46 LTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQ 105
Db         23 LARKMRSDVTDLLDIYVERQG-----LDASISVAAVDGVPTAAV--ERWAEQTGTQRLLD 75

Qy        106 N---YEAYSHLLCYLRGLNRQA---ATAELRRSLA-------HFCTSLQGLLGSIAGVMA 152
Db         76 NLAAYRAFRTLLAQMLEEQRELLGDTDAELGPALAAMLLQVSAFVYHLEELL-----ELE 130

Qy        153 ALGYPLPQPLPGTEPTWTPGPAH-SDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQP 211
Db        131 SRGAPAEE---GSEP---PAPPRLSLFEQKLRGLRVLRELAQWAVRSVRDLRQLSKHGPG 184

Qy        212 PAAAVTL 218
Db        185 SGAALGL 191


RESULT 7
T38943
probable guanine nucleotide binding protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38943
R;Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z21819
A;Accession: T38943
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1313 <SKE>
A;Cross-references: EMBL:Z95396; PIDN:CAB08769.1; GSPDB:GN00066; SPDB:SPAC57A7.11
A;Experimental source: strain 972h-; cosmid c57A7
C;Genetics:
A;Gene: SPDB:SPAC57A7.11
A;Map position: 1

  Query Match             7.3%;  Score 89.5;  DB 2;  Length 1313;
  Best Local Similarity   22.8%;  Pred. No. 8;
  Matches   61;  Conservative   32;  Mismatches   69;  Indels  105;  Gaps   14;

Qy         15 LCTVLWHLPAV--PALNR-------TGDPGPGPSI---------QKTYDLTRYLEHQLRS 56
Db        811 LAFLLQHLPALHKASLSKDTDTNSVTSDPKPHPFVPSVSENKILNRSFSLTRSLKGLALS 870

Qy         57 LAG------------------TYLNYL------GPP-FNEPDFNPPRLGAETLPRATVDL 91
Db        871 LAGSDRASELLSLNGENKPAESNLNHLTSAKVPGPPAFNELEY----------QSELDM 919

Qy         92 EVWRSLND---------KLRLTQNYEAYSHLLCYLRGL---NRQAATAELRRSLAHFCTS 139
Db        920 PLTSYLFDWSRKYFTEPQMRPNEDDEPGS--ICYNQRLWRRNRNEKLIYRTRPLAEYSTN 977

Qy        140 ------LQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKE--- 190
Db        978 GRWNQQLMTPNNTIA----------------------PRKLMFHQFEDQLITLGDKDI 1013

Qy        191 LQTWLWRSAKDFNRLKKKMQPPAAAVT 217
Db       1014 IQVWDWRR----NRCLNSFKTSASATT 1036


RESULT 8
D83588
conserved hypothetical protein PA0454 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83588
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: D83588
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-733 <STO>
A;Cross-references: GB:AE004483; GB:AE004091; NID:g9946313; PIDN:AAG03843.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0454
C;Superfamily: hypothetical protein HI1680

  Query Match             7.3%;  Score 89;  DB 2;  Length 733;
  Best Local Similarity   26.6%;  Pred. No. 4.3;
  Matches   58;  Conservative   26;  Mismatches   66;  Indels   68;  Gaps   13;

Qy         28 LNRTGDPGPGPSIQKTYDLTRYLE--------HQLRSLAGTYLNYLGPPFNEPD--FNPP 77
Db        230 LNRLGHGRPGPKV------SRYLKLYFIAQDVHERASSSHYPYNRLAEAFFHSDVLFRCQ 283

Qy         78 RL------GAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRR 131
Db        284 RLLNQQGKACQALARA-------------IRLRQPFD-YA---------DRELALEDLQA 320

Qy        132 SLAHF----CTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL 187
Db        321 SLEHLRQQSNPAWKGLLRSL-GALAANLTTLDRKLAGA--------SNPDAIADEQDSAL 371

Qy        188 LKELQTWLWRSAKD-FNRLKKKMQPPAAAVTLHLGAHG 224
Db        372 LDRSP----RSLKDAFERLRQQLTP-----TSLLFRHG 400


RESULT 9
T00247
zinc finger protein wiz - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 05-Nov-1999
C;Accession: T00247
R;Matsumoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Wanaka, A.; Tohyama, M.
submitted to the EMBL Data Library, March 1998
A;Description: Molecular cloning and distinct developmental expression pattern of splice
A;Reference number: Z14130
A;Accession: T00247
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
```

A;Residues: 1-955 <MAT>
A;Cross-references: EMBL:AB012266; NID:d1227741; PIDN:BAA32791.1; PID:d1033757
A;Experimental source: brain
C;Genetics:
A;Gene: wiz

Query Match 7.1%; Score 87; DB 2; Length 955;
Best Local Similarity 27.8%; Pred. No. 9.2;
Matches 35; Conservative 19; Mismatches 48; Indels 24; Gaps 7;

```
Qy         22 LPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAG---TYLNYLGP-----PFNEPD 73
              ||  |  :| | || ||:      : |  |  | : |   :  :||||     |  |
Db        615 LPLSPLASRPGKPGAGPT-----QVPR--ELSLSPITGSKPSAASYLGPVATKRPLQEDR 667

Qy         74 FNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL---NRQAATAELR 130
              | |  : |:|  :  :  :  ::|::|        : :   | | ||   ||:|  :  |
Db        668 FLPAEVKAKTYIQTELPFKA-KTLHEK-----TSHSSTEACCELCGLYFENRKALASHAR 721

Qy        131 RSLAHF 136
                |  |
Db        722 AHLRQF 727
```

RESULT 10
T00248
zinc finger protein wizL - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 05-Nov-1999
C;Accession: T00248
R;Matsumoto, K.; Ishii, N.; Yoshida, S.; Shiosaka, S.; Wanaka, A.; Tohyama, M.
submitted to the EMBL Data Library, March 1998
A;Description: Molecular cloning and distinct developmental expression pattern of splice
A;Reference number: Z14130
A;Accession: T00248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1561 <MAT>
A;Cross-references: EMBL:AB012265; NID:d1227740; PIDN:BAA32790.1; PID:d1033756
A;Experimental source: brain
C;Genetics:
A;Gene: wiz

Query Match 7.1%; Score 87; DB 2; Length 1561;
Best Local Similarity 27.8%; Pred. No. 17;
Matches 35; Conservative 19; Mismatches 48; Indels 24; Gaps 7;

```
Qy         22 LPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAG---TYLNYLGP-----PFNEPD 73
              ||  |  :| | || ||:      : |  |  |  : |   :  :||||     |  |
Db       1221 LPLSPLASRPGKPGAGPT-----QVPR--ELSLSPITGSKPSAASYLGPVATKRPLQEDR 1273

Qy         74 FNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL---NRQAATAELR 130
              | |  : |:|  :  :  :  ::|::|        : :   | | ||   ||:|  :  |
Db       1274 FLPAEVKAKTYIQTELPFKA-KTLHEK-----TSHSSTEACCELCGLYFENRKALASHAR 1327

Qy        131 RSLAHF 136
                |  |
Db       1328 AHLRQF 1333
```

RESULT 11
T34916
transferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 02-Sep-2000
C;Accession: T34916
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z21558
A;Accession: T34916
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-640 <OLI>
A;Cross-references: EMBL:AL021409; PIDN:CAA16181.1; GSPDB:GN00070; SCOEDB:SC3F7.10
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC3F7.10
C;Superfamily: glycine C-acetyltransferase homology
F;287-624/Domain: glycine C-acetyltransferase homology <GCA>

Query Match 7.1%; Score 86.5; DB 2; Length 640;
Best Local Similarity 27.2%; Pred. No. 6.1;
Matches 50; Conservative 13; Mismatches 64; Indels 57; Gaps 9;

```
Qy         22 LPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQL---RSLAGTYLNYLGPPFNEPDFNPPR 78
              ||||||   ||||     :      |:|  |||     | |
Db        100 LPAVPA-EGTGDPVTEAVVAAAMRQTQYQRHQLGLDADLEG-----------------E 140

Qy         79 LGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCT 138
              || :::   :|  |       ::| ||                |    || |   |:||
Db        141 LGVDSVVLTSVVAEA----TERLGLT--------------GAAPDAAGATTLRALA---D 179

Qy        139 SLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRS 198
              :|:||       | || |  :|:  | |     | |  |                |  ||
Db        180 ALRGL------VAAAPGTAVPEAAPATGAA-APAPGRSGNAP-------APGADGWDHRS 225

Qy        199 AKDF 202
                |||
Db        226 MKDF 229
```

RESULT 12
C38604
poly(3-hydroxyalkanoate) polymerase (EC 2.7.7.-) 3 - Pseudomonas oleovorans
C;Species: Pseudomonas oleovorans
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 08-Sep-2000
C;Accession: C38604
R;Huisman, G.W.; Wonink, E.; Meima, R.; Kazemier, B.; Terpstra, P.; Witholt, B.
J. Biol. Chem. 266, 2191-2198, 1991
A;Title: Metabolism of poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas oleovorans. Ident
A;Reference number: A38604; MUID:91115830; PMID:1989978
A;Accession: C38604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-560 <HUI>
A;Cross-references: GB:M58445; NID:g151441; PIDN:AAA25934.1; PID:g151444
C;Superfamily: poly(3-hydroxyalkanoic acid) synthase phbC
C;Keywords: nucleotidyltransferase

Query Match 7.0%; Score 85.5; DB 2; Length 560;
Best Local Similarity 30.2%; Pred. No. 6.4;
Matches 38; Conservative 17; Mismatches 42; Indels 29; Gaps 7;

```
Qy         77 PRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRRSL--A 134
              |  |  |||  ::::      |  | |           |:  || ::||:    ||  |  |
Db          5 PAKGTPTLPATSMNVQ-----NAILGLRGR-----DLISTLRNVSRQS----LRHPLHTA 50

Qy        135 HFCTSLQGLLGSIAGVMAALGYPLPQPLPG----TEPTWTPGPAHSDFLQKMDDFWLLKE 190
              |   :| | || :      ||     || |      ::|||:  | :   ||     :|  :
Db         51 HHLLALGGQLGRV-----ILGDTPLQPNPRDPRFSDPTWSQNPFYRRGLQA----YLAWQ 101

Qy        191 LQTWLW 196
                || ||
Db        102 KQTRLW 107
```

RESULT 13
F72511
probable cystathionine gamma-synthase APE2068 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: F72511
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: F72511
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-389 <KAW>
A;Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81078.1; PID:g5105766
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2068
C;Superfamily: O-succinylhomoserine (thiol)-lyase

Query Match 6.9%; Score 84.5; DB 2; Length 389;
Best Local Similarity 22.9%; Pred. No. 5;
Matches 47; Conservative 19; Mismatches 78; Indels 61; Gaps 8;

```
Qy          43 TYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSL----- 97
               ||  || |   | |: |  :   |||: :                 :||  |  |
Db         107 TYGSTRSLLEMLSSITGIEVRLAGPPWED---------------LLDLVCWADLIIVES 150

Qy          98 --NDKLRLTQNYEAYSHLLCYLRGLNRQAATAELR-------------RSL---AHFCTS 139
                 |  ||:              | |: |:| :  :|             | |   ||:  |
Db         151 MANPTLRVPP-----------LSGIYREAGSCGVRVVVDNTFATPIAYRPLERGAHY--S 197

Qy         140 LQGLLGSIAGVMAALGYPLPQPL-PGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRS 198
               |: |   |||    :|  |   :    || |         :| :| :          ||
Db         198 LESLTKYIAGHNDVVGGSLSGRVEEDLEPLWNMRKILGTIMQPIDAY--------LAWRG 249

Qy         199 AKDFNRLKKKMQPPAAAVTLHLGAH 223
               |     :    |  |   | :|
Db         250 MKTLKARFEAQSRAAVEVAEWLESH 274
```

RESULT 14
AD3633
hypothetical protein BMEII0989 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AD3633
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AD3633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-200 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL54231.1; PID:g17985203; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0989
A;Map position: II

Query Match 6.9%; Score 84; DB 2; Length 200;
Best Local Similarity 24.3%; Pred. No. 2.4;
Matches 56; Conservative 22; Mismatches 84; Indels 68; Gaps 13;

```
Qy          10 GMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNYLGPPF 69
               |:||    |   ||::||  :|       |:    :    ||   ||  |  :   :
Db           2 GLLAGAALV---LPSLPAEAKTQQAAMPPNATSPHQADVYL---LRGFADIFSTGI---- 51

Qy          70 NEPDFNPPRLGAETLPRATVDLEV-----WRSLNDKLRLTQNYEAYSHLLCYLRG--LNR 122
                        :||| |  | |:  |      || : :  |:  : :   ||   | |  |
Db          52 -------DEIGAE-LQAAGVNAHVQGHAAWRLVLN--RIVADQQKNGHLPVVLIGHSLGA 101

Qy         123 QAA---TAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGT----------EPTW 169
                 ||     || |         :|:        || |   | ||||          :  |
Db         102 NAAIYIAEELER---------RGIAVDYMATFAATG---PDPLPGNVRRVVNFYFKQHGW 149

Qy         170 ----TPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAA 215
                   |||     |:  |            : :|||      : | |  |
Db         150 GLPLVPGPRFHGHLENRD-----------FSNAKDVGHFNIEKQRPLQA 187
```

RESULT 15
F75439
probable cell wall synthesis protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: F75439
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: F75439
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <WHI>
A;Cross-references: GB:AE001958; GB:AE000513; NID:g6458805; PIDN:AAF10649.1; PID:g645880
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1076
A;Map position: 1

Query Match 6.9%; Score 84; DB 2; Length 411;
Best Local Similarity 24.6%; Pred. No. 6;
Matches 56; Conservative 25; Mismatches 77; Indels 70; Gaps 11;

```
Qy           2 DLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTY 61
               :: | | | : | | : :   | || |  :|  | |          ||     :| :| |
Db         212 EVLAADRWALRAELFSEMSLRPGVPLLLLSGG-GRGHYAAAADVLT-----ELGNL-GRA 264

Qy          62 LNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLN 121
               :  | |   : :       | ||:  |||                      | | : | |
Db         265 VQVLVPASRQGE------GTETIGGATV----------------------HHLGFRRDLP 296

Qy         122 RQAATAELRRSLAHFCTSLQGLLGSIAGV----MAALGYPLP--QPLPGTEPTWTPGPAH 175
               |  | ::|             ::|   |:      ||| ||    |:|| |        :
Db         297 RLLAASDL-------------VVGKAGGLTVAEATALGVPLVIYAPIPGQE------EHN 337

Qy         176 SDFLQKMDDFWLLKELQTWLW-RSAKDFNRLKKKMQPPAAAVTLHLGA 222
               :|||::             || |:  |   |  :  ||    |  ||
Db         338 ADFLERHG---------AGLWARARHDVRPLVLRALDPAEHARLSAGA 376
```

Search completed: August 10, 2004, 06:44:11
Job time : 16.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:43:28 ; Search time 8 Seconds
(without alignments)
1464.473 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	218.5	17.8	204	1	CTF2_MOUSE	P83714 mus musculu
2	124.5	10.2	203	1	CTF1_RAT	Q63086 rattus norv
3	118.5	9.7	203	1	CTF1_MOUSE	Q60753 mus musculu
4	96.5	7.9	455	1	CTY2_MOUSE	Q9cwv7 mus musculu
5	94	7.7	423	1	CST_MOUSE	Q9jhe4 m galactosy
6	93	7.6	619	1	NXF1_HUMAN	Q9ubu9 homo sapien
7	91.5	7.5	201	1	CTF1_HUMAN	Q16619 homo sapien
8	91.5	7.5	1182	1	HAIR_MOUSE	Q61645 mus musculu
9	91	7.4	195	1	CNTF_CHICK	Q02011 gallus gall
10	89.5	7.3	1313	1	MIP1_SCHPO	P87141 schizosacch
11	87.5	7.1	1181	1	HAIR_RAT	P97609 rattus norv
12	86	7.0	1009	1	M2B2_HUMAN	Q9y2e5 homo sapien
13	85.5	7.0	560	1	PHAC_PSEOL	P26496 pseudomonas
14	84	6.9	200	1	CNTF_PIG	O02732 sus scrofa
15	83	6.8	452	1	TLL_DROME	P18102 drosophila
16	82.5	6.7	1189	1	HAIR_HUMAN	O43593 homo sapien
17	82	6.7	724	1	P85B_BOVIN	P23726 bos taurus
18	81	6.6	291	1	YM32_MYCTU	Q10515 mycobacteri
19	81	6.6	618	1	NXF1_RAT	O88984 rattus norv
20	80.5	6.6	294	1	HUPK_RHOCA	P30797 rhodobacter
21	80.5	6.6	870	1	BCA1_HUMAN	P56945 homo sapien
22	80.5	6.6	1473	1	NAL1_HUMAN	Q9c000 homo sapien
23	80	6.5	200	1	CNTF_RAT	P20294 rattus norv
24	80	6.5	422	1	Y140_HUMAN	Q14153 homo sapien
25	80	6.5	2261	1	RRPL_MUMPM	P30929 mumps virus
26	79	6.4	552	1	MPP2_MOUSE	Q9wv34 mus musculu
27	79	6.4	1621	1	ALK_MOUSE	P97793 mus musculu
28	78.5	6.4	390	1	YL28_STRCO	P40181 streptomyce
29	78.5	6.4	1102	1	CARB_STRCO	Q9kxr6 streptomyce
30	78	6.4	3680	1	DMD_CANFA	O97592 canis famil
31	77	6.3	586	1	UL84_HCMVA	P16727 human cytom
32	77	6.3	587	1	UL84_HCMVT	P29839 human cytom
33	76.5	6.2	571	1	ATKA_MYCTU	P96371 mycobacteri
34	76.5	6.2	572	1	SYM_AERPE	Q9ycy3 aeropyrum p
35	76.5	6.2	995	1	M2B2_PIG	Q28949 sus scrofa
36	76	6.2	296	1	RECO_ANASP	Q8yp19 anabaena sp
37	76	6.2	416	1	TGAS_STRCJ	Q8gr90 streptomyce
38	76	6.2	1001	1	PTPX_MOUSE	P80560 mus musculu
39	76	6.2	1132	1	BAT3_HUMAN	P46379 homo sapien
40	75.5	6.2	830	1	VPP3_HUMAN	Q13488 h vacuolar
41	75	6.1	199	1	CNTF_RABIT	P14188 oryctolagus
42	75	6.1	315	1	YNEF_ECOLI	P76147 escherichia
43	75	6.1	343	1	DFRA_SYNY3	P73212 synechocyst
44	75	6.1	450	1	TLL_DROVI	O16845 drosophila
45	75	6.1	917	1	SYI_STAAM	Q99us4 staphylococ

ALIGNMENTS

RESULT 1
CTF2_MOUSE
ID CTF2_MOUSE STANDARD; PRT; 204 AA.
AC P83714;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cardiotrophin-2 precursor (CT-2).
GN CTF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hasegawa M., Habu K., Adachi Y., Natori O.;
RT "Cardiotrophin-2 increases the platelet counts with splenomegaly in
RT vivo.";
RL Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Increases the platelet count associated with
CC splenomegaly.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC ---

CC ---
DR EMBL; AB125661; BAD01485.1; -.
KW Cytokine; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 204 CARDIOTROPHIN-2.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 204 AA; 22000 MW; B05566C1E7C3D9C0 CRC64;

Query Match 17.8%; Score 218.5; DB 1; Length 204;
Best Local Similarity 30.3%; Pred. No. 2.8e-13;
Matches 64; Conservative 31; Mismatches 81; Indels 35; Gaps 6;

```
Qy         14 CLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPD 73
              || ::|     :| |:       |   | : | |  |::    :|  ||| : | ||::|
Db         10 CLLSLL-----LPPLSPAAPISPSEPIGQAYSLALYMQKNTSALLQTYLQHQGSPFSDPG 64

Qy         74 FNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLR--------------G 119
              |: | |   ||| | |  : | :: |  ||::   |:  |  :|:
Db         65 FSAPELQLSTLPSAAVSFKTWHAMEDAERLSRAQGAFLALTQHLQLVGDDQSYLNPGSPI 124

Qy        120 LNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPT--WTPGPAHSD 177
              | |  | ||          |||||::| :| ||| |:|       | |  : |  | |
Db        125 LLAQLGAARLR---------AQGLLGNMAAIMTALGLPIPP----EEDTLGFVPFGA-SA 170

Qy        178 FLQKMDDFWLLKELQTWLWRSAKDFNRLKKK 208
```

```
Db        171 FERKCRGYIVTREYGHWTDRAVRDLALLKAK 201


RESULT 2
CTF1_RAT
ID   CTF1_RAT       STANDARD;      PRT;   203 AA.
AC   Q63086;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Cardiotrophin-1 (CT-1).
GN   CTF1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Heart;
RX   MEDLINE=96193659; PubMed=8604995;
RA   Ishikawa M., Saito Y., Miyamoto Y., Kuwahara K., Ogawa E.,
RA   Nakagawa O., Harada M., Masuda I., Nakao K.;
RT   "cDNA cloning of rat cardiotrophin-1 (CT-1): augmented expression of
RT   CT-1 gene in ventricle of genetically hypertensive rats.";
RL   Biochem. Biophys. Res. Commun. 219:377-381(1996).
CC   -!- FUNCTION: Induces cardiac myocyte hypertrophy in vitro. Binds to
CC       and activates the ILST/gp130 receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in the ventricle and atrium of adult
CC       rats. Also detected in the lung, kidney, liver, skeletal muscle,
CC       stomach and urinary bladder. Not detected in brain, colon, testis,
CC       spleen or thymus. Overexpressed in the ventricules in the case of
CC       hypertension and hypertrophy.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D78591; BAA11427.1; -.
DR   PIR; JC4645; JC4645.
KW   Cytokine.
SQ   SEQUENCE   203 AA;  21439 MW;  DF8B921A2FA1C832 CRC64;

  Query Match             10.2%;  Score 124.5;  DB 1;  Length 203;
  Best Local Similarity   28.1%;  Pred. No. 0.00018;
  Matches   50;  Conservative   24;  Mismatches   85;  Indels   19;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |      :   | |
Db         27 IRQTHNLARLLTKYADQLLEEYVQQQGEPPGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||       :: |  ::  |:
Db         84 ---SERLRQDAAALSALPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGY----PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||      |:|:|:  |  :|  |   |  |:   :       |: |:  |  :|
Db        141 AALGAAARGPVPEPV-ATSALFTSNSAAGVFSAKVLGLHVCGLYGEWVSRTEGDLGQL 197


RESULT 3
CTF1_MOUSE
ID   CTF1_MOUSE     STANDARD;      PRT;   203 AA.
AC   Q60753;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Cardiotrophin-1 (CT-1).
GN   CTF1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95166785; PubMed=7862649;
RA   Pennica D., King K.L., Shaw K.J., Luis E., Rullamas J., Luoh S.-M.,
RA   Darbonne W.C., Knutzon D.S., Yen R., Chien K.R., Baker J.B.,
RA   Wood W.I.;
RT   "Expression cloning of cardiotrophin 1, a cytokine that induces
RT   cardiac myocyte hypertrophy.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:1142-1146(1995).
CC   -!- FUNCTION: Induces cardiac myocyte hypertrophy in vitro. Binds to
CC       and activates the ILST/gp130 receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Highly expressed in heart, skeletal muscle,
CC       liver, lung and kidney. Lower levels in testis and brain. No
CC       expression in spleen.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U18366; AAC52173.1; -.
DR   PIR; I49153; I49153.
DR   MGD; MGI:105115; Ctf1.
KW   Cytokine.
SQ   SEQUENCE   203 AA;  21509 MW;  8B3D414A0B3B232F CRC64;

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 0.00064;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |      :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPPGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||       :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||      | | | |  :|       |  |:  | :       |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 4
CTY2_MOUSE
ID   CTY2_MOUSE     STANDARD;      PRT;   455 AA.
AC   Q9CWV7; Q8R163;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein C20orf162 homolog.
GN   C20ORF162.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryonic stem cells;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
```

```
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain, and Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 1 SWIM-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK010358; BAB26878.1; -.
DR   EMBL; BC025184; AAH25184.1; ALT_INIT.
DR   EMBL; BC046831; AAH46831.1; -.
DR   MGD; MGI:1919221; 2410003H12Rik.
DR   InterPro; IPR007527; Znf_SWIM.
DR   Pfam; PF04434; SWIM; 1.
KW   Zinc-finger.
FT   ZN_FING     342    375       SWIM-TYPE.
SQ   SEQUENCE   455 AA;  51481 MW;  201886B814EB9CFF CRC64;

  Query Match             7.9%;  Score 96.5;  DB 1;  Length 455;
  Best Local Similarity   22.4%;  Pred. No. 0.21;
  Matches   59;  Conservative   24;  Mismatches   89;  Indels   91;  Gaps   12;

Qy          7 DSWGMLACLCTVLW-HLPAVPALNRTGDPG---------PGPSIQKTYDLTRYLEHQLRS 56
              |::   :|    |: | | |  ::|| :|          || :|    | | :
Db         29 DAFNYQSCFMQDLFAHFPEVLFIHRTYNPRGKVLYTFLVDGPRVQVEGPLARAV------ 82

Qy         57 LAGTYLNYLGPPFNE------------PDFNPPRLGAETLPRATVDLEVWRSLNDKLRLT 104
                |   | ||                |||                 |  :|  |
Db         83 -------YFAIPTNEDARGLAQMFQVFKKFNP---------------AWERVNTIL--- 116

Qy        105 QNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQP--- 161
                  |  |  |  |  :  |||:  |  | |  |||    :          || ||
Db        117 ----VDPHFL-LLPTLTMEFPTAEVLLSAFHICKFLQGKFYQL---------PLEQPVQR 162

Qy        162 -----LPGTEPTWTPGPAH-----------SDFLQKMDDFWLLKELQTWL---WRSAKDF 202
                  |  |  :  |  |              |  |  ::   |||  :  :  ||   |||
Db        163 LLLSSLQSTMCSATAGNLRKLYTLLNNCIPSSRLPELHSHWLLND-RIWLAHRWRSRAQS 221

Qy        203 NRLKKKMQPPAAAVTLHLGAHGF 225
                :|  : ::  |  ::   |  |
Db        222 SRYFQSLEIMAHILSQFFGTTPF 244


RESULT 5
CST_MOUSE
ID   CST_MOUSE      STANDARD;      PRT;   423 AA.
AC   Q9JHE4; Q9D8V6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Galactosylceramide sulfotransferase (EC 2.8.2.11) (GalCer
DE   sulfotransferase) (Cerebroside sulfotransferase) (3'-
DE   phosphoadenylylsulfate:galactosylceramide 3'-sulfotransferase)
DE   (3'-phosphoadenosine-5'phosphosulfate:GalCer sulfotransferase).
GN   CST OR GCST.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Kidney;
RX   MEDLINE=20193614; PubMed=10727929;
RA   Hirahara Y., Tsuda M., Wada Y., Honke K.;
RT   "cDNA cloning, genomic cloning, and tissue-specific regulation of
RT   mouse cerebroside sulfotransferase.";
RL   Eur. J. Biochem. 267:1909-1917(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Pancreas;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
```

```
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=21927584; PubMed=11917099;
RA   Honke K., Hirahara Y., Dupree J., Suzuki K., Popko B., Fukushima K.,
RA   Fukushima J., Nagasawa T., Yoshida N., Wada Y., Taniguchi N.;
RT   "Paranodal junction formation and spermatogenesis require
RT   sulfoglycolipids.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4227-4232(2002).
CC   -!- FUNCTION: Catalyzes the sulfation of membrane glycolipids. Seems
CC       to prefer beta-glycosides at the nonreducing termini of sugar
CC       chains attached to a lipid moiety. Catalyzes the synthesis of
CC       HSO3-3-galactosylceramide (sulfatide), a major lipid component of
CC       the myelin sheath and of HSO3-3-monogalactosylalkylacylglycerol
CC       (seminolipid), present in spermatocytes. Also acts on
CC       lactosylceramide, galactosyl 1-alkyl-2-sn-glycerol and galactosyl
CC       diacylglycerol (in vitro).
CC   -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + a
CC       galactosylceramide = adenosine 3',5'-bisphosphate +
CC       galactosylceramide sulfate.
CC   -!- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate +
CC       monogalactosylalkylacylglycerol = adenosine 3',5'-bisphosphate +
CC       monogalactosylalkylacylglycerol sulfate.
CC   -!- PATHWAY: Sphingolipid and glycerolipid biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi membrane (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in brain, testis, kidney, stomach,
CC       small intestine, liver, and lung. Not detected in heart, skeletal
CC       muscle, and spleen.
CC   -!- MISCELLANEOUS: Mice homozygous for a null mutation of the CST gene
CC       display hindlimb weakness from week 6 of age and subsequently show
CC       a prenounced tremor and progressive ataxia. Myelin vacuaolation is
CC       observed in the cerebellar white matter, diencephalon, brainstem
CC       and spinal anterior column. Male mice were infertile due to a
CC       blocked spermatogenesis.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB032940; BAA93009.1; -.
DR   EMBL; AB032939; BAA93008.1; -.
DR   EMBL; AK007645; BAB25160.1; -.
DR   EMBL; BC026806; AAH26806.1; -.
DR   MGD; MGI:1858277; Gcst.
DR   GO; GO:0016021; C:integral to membrane; IC.
DR   GO; GO:0001733; F:galactosylceramide sulfotransferase activity; IDA.
DR   GO; GO:0006682; P:galactosylceramide biosynthesis; IDA.
DR   GO; GO:0042552; P:myelination; IMP.
DR   GO; GO:0007283; P:spermatogenesis; IMP.
KW   Transferase; Transmembrane; Glycoprotein.
FT   DOMAIN        1     12       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     13     35       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       36    423       LUMENAL, CATALYTIC (POTENTIAL).
FT   CARBOHYD     66     66       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    312    312       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     18     18       L -> P (IN REF. 2).
FT   CONFLICT    263    263       R -> Q (IN REF. 2).
FT   CONFLICT    271    271       E -> K (IN REF. 2).
FT   CONFLICT    358    358       Q -> R (IN REF. 1).
FT   CONFLICT    392    392       R -> G (IN REF. 2).
FT   CONFLICT    398    398       I -> T (IN REF. 1).
SQ   SEQUENCE   423 AA;  48968 MW;  FD54A1A71F4AEE46 CRC64;

  Query Match             7.7%;  Score 94;  DB 1;  Length 423;
  Best Local Similarity   22.7%;  Pred. No. 0.32;
  Matches   55;  Conservative   23;  Mismatches   60;  Indels  104;  Gaps   12;

Qy         44 YDLTRYLEHQLRSLAGTYLNY---LGPP----------------------FNEP----- 72
              || : |  | ||:|     | |    | |                      |:|
Db        204 YDPSSYNAHYLRNLLFFDLGYDSSLDPASPRVQEHILEVERRFHLVLLQEYFDESLVLLR 263

Qy         73 ------------------DFNPPRLGAETLPRAT----VDLEVWRSLNDKLRLTQNYEA 109
                               |   |||  |   |||    :|: ::|  |      :  ||
Db        264 ELLCWDLEDVLYFKLNARRDSPVPRLSGELYRRATAWNLLDVRLYRHFN--ASFWRKVEA 321

Qy        110 YSHLLCYLRGLNRQA-ATAELRRS---LAHFCTSLQGLLGSIA-----------GVMAA 153
              :     |  | |   ||||::   : | |      :|: |             |: :
Db        322 F--------GRERMAREVAELRQANEHMRHICIDGGQAVGAEAIQDSAMQPWQPLGIKSI 373

Qy        154 LGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQ----------TWLWRSAKDF 202
              ||| | : :          || |    ::|    | |:|          | ||:  :||
Db        374 LGYNLKKSI---------GPQHEQLCRRM----LTPEIQYLSDLGANLWVTKLWKFLRDF 420

Qy        203 NR 204
                |
Db        421 LR 422


RESULT 6
NXF1_HUMAN
ID   NXF1_HUMAN     STANDARD;      PRT;   619 AA.
AC   Q9UBU9; Q99799; Q9UQL2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Nuclear RNA export factor 1 (Tip associating protein) (Tip-associated
DE   protein) (mRNA export factor TAP).
GN   NXF1 OR TAP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS.
RC   TISSUE=Cervical carcinoma;
RX   MEDLINE=99219873; PubMed=10202158;
RA   Braun I.C., Rohrbach E., Schmitt C., Izaurralde E.;
RT   "TAP binds to the constitutive transport element (CTE) through a novel
RT   RNA-binding motif that is sufficient to promote CTE-dependent RNA
RT   export from the nucleus.";
RL   EMBO J. 18:1953-1965(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99257272; PubMed=10323864;
RA   Kang Y., Cullen B.R.;
RT   "The human Tap protein is a nuclear mRNA export factor that contains
RT   novel RNA-binding and nucleocytoplasmic transport sequences.";
RL   Genes Dev. 13:1126-1139(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99384298; PubMed=10454577;
RA   Bear J., Tan W., Zolotukhin A.S., Tabernero C., Hudson E.A.,
RA   Felber B.K.;
RT   "Identification of novel import and export signals of human TAP, the
RT   protein that binds to the constitutive transport element of the type
RT   D retrovirus mRNAs.";
RL   Mol. Cell. Biol. 19:6306-6317(1999).
RN   [4]
```

RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP SEQUENCE OF 61-619 FROM N.A.
RC TISSUE=Lymphocytes;
RX MEDLINE=97318898; PubMed=9175835;
RA Yoon D.-W., Lee H., Seol W., DeMaria M., Rosenzweig M., Jung J.U.;
RT "Tap: a novel cellular protein that interacts with tip of herpesvirus
RT saimiri and induces lymphocyte aggregation.";
RL Immunity 6:571-582(1997).
RN [7]
RP FUNCTION.
RX MEDLINE=98325379; PubMed=9660949;
RA Grueter P., Tabernero C., von Kobbe C., Schmitt C., Saavedra C.,
RA Bachi A., Wilm M., Felber B.K., Izaurralde E.;
RT "TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export
RT from the nucleus.";
RL Mol. Cell 1:649-659(1998).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=21282872; PubMed=11259411;
RA Braun I.C., Herold A., Rode M., Conti E., Izaurralde E.;
RT "Overexpression of TAP/p15 heterodimers bypasses nuclear retention and
RT stimulates nuclear mRNA export.";
RL J. Biol. Chem. 276:20536-20543(2001).
RN [9]
RP CHARACTERIZATION.
RX MEDLINE=20132240; PubMed=10668806;
RA Bachi A., Braun I.C., Rodrigues J.P., Pante N., Ribbeck K.,
RA von Kobbe C., Kutay U., Wilm M., Gorlich D., Carmo-Fonseca M.,
RA Izaurralde E.;
RT "The C-terminal domain of TAP interacts with the nuclear pore complex
RT and promotes export of specific CTE-bearing RNA substrates.";
RL RNA 6:136-158(2000).
RN [10]
RP MUTAGENESIS.
RX MEDLINE=21151125; PubMed=11256625;
RA Suyama M., Doerks T., Braun I.C., Sattler M., Izaurralde E., Bork P.;
RT "Prediction of structural domains of TAP reveals details of its
RT interaction with p15 and nucleoporins.";
RL EMBO Rep. 1:53-58(2000).
RN [11]
RP INTERACTION WITH THOC4 AND THE EXON JUNCTION COMPLEX.
RX MEDLINE=21564074; PubMed=11707413;
RA Kataoka N., Diem M.D., Kim V.N., Yong J., Dreyfuss G.;
RT "Magoh, a human homolog of Drosophila mago nashi protein, is a
RT component of the splicing-dependent exon-exon junction complex.";
RL EMBO J. 20:6424-6433(2001).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 102-372.
RX MEDLINE=20514125; PubMed=11060011;
RA Liker E., Fernandez E., Izaurralde E., Conti E.;
RT "The structure of the mRNA export factor TAP reveals a cis arrangement
RT of a non-canonical RNP domain and an LRR domain.";
RL EMBO J. 19:5587-5598(2000).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF COMPLEX WITH NXT1, AND X-RAY
RP CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF COMPLEX WITH NXT1-FG-REPEAT.
RX MEDLINE=21468398; PubMed=11583626;
RA Fribourg S., Braun I.C., Izaurralde E., Conti E.;
RT "Structural basis for the recognition of a nucleoporin FG repeat by
RT the NTF2-like domain of the TAP/p15 mRNA nuclear export factor.";
RL Mol. Cell 8:645-656(2001).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (1.0 ANGSTROMS) OF 561-619 IN COMPLEX WITH
RP RANBP3.
RX MEDLINE=22469929; PubMed=12581645;
RA Grant R.P., Neuhaus D., Stewart M.;
RT "Structural basis for the interaction between the Tap/NXF1 UBA domain
RT and FG nucleoporins at 1A resolution.";
RL J. Mol. Biol. 326:849-858(2003).
RN [15]
RP STRUCTURE BY NMR OF 551-619, AND MUTAGENESIS OF PHE-617.
RX MEDLINE=21912422; PubMed=11875519;
RA Grant R.P., Hurt E., Neuhaus D., Stewart M.;
RT "Structure of the C-terminal FG-nucleoporin binding domain of
RT Tap/NXF1.";
RL Nat. Struct. Biol. 9:247-251(2002).
CC -!- FUNCTION: Involved in the nuclear export of mRNA species bearing
CC retroviral constitutive transport elements (CTE) and in the export
CC of mRNA from the nucleus to the cytoplasm.
CC -!- SUBUNIT: Interacts with NXT1, NXT2, E1B-AP5, RAE1, THOC4
CC and with several nucleoporins. Is part of the exon
CC junction complex (EJC) containing NCBP1, NCBP2, RNPS1, RBM8A,
CC SRRM1, NXF1, NFX1, UPF3B, UPF2 and THOC4.
CC -!- SUBCELLULAR LOCATION: Nuclear; localized predominantly in the
CC nucleoplasm and at both the nucleoplasmic and cytoplasmic faces of
CC the nuclear pore complex. Shuttles between the nucleus and the
CC cytoplasm.
CC -!- TISSUE SPECIFICITY: Expressed ubiquitously.
CC -!- DOMAIN: The minimal CTE binding domain consists of an RNP-
CC type RNA binding domain (RBD) and leucine-rich repeats.
CC -!- DOMAIN: The nucleoporin binding domain consists of a NTF2-like
CC domain and a UBA-like domain. The NTF2 domain heterodimerizes with
CC NXT1 AND NXT2. The formation of NXF1/NXT1 heterodimers is required
CC for NXF1-mediated nuclear mRNA export. The UBA-like domain
CC mediates direct interactions with nucleoporin-FG-repeats and is
CC necessary and sufficient for localization of NXF1 to the nuclear
CC rim. The conserved loop 594-NWD-596 of the UBA domain has a
CC critical role in the interaction with nucleoporins.
CC -!- DOMAIN: The leucine-rich repeats and the NTF2-domain are
CC essential for the export of mRNA from the nucleus.
CC -!- MISCELLANEOUS: The RNA-binding domain is a non-canonical RNP-type
CC domain.
CC -!- SIMILARITY: Belongs to the NXF family.
CC -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 NTF2 domain.
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
CC -!- SIMILARITY: Contains 1 UBA-like domain.
CC --

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ132712; CAA10753.1; -.
DR   EMBL; AF112880; AAD39102.1; -.
DR   EMBL; AF126246; AAD20016.1; -.
DR   EMBL; AK027192; -; NOT_ANNOTATED_CDS.
DR   EMBL; BC004904; AAH04904.1; -.
DR   EMBL; BC028041; AAH28041.1; -.
DR   EMBL; U80073; AAB81111.1; -.
DR   PDB; 1FO1; 03-NOV-00.
DR   PDB; 1FT8; 11-DEC-00.
DR   PDB; 1KOH; 27-FEB-02.
DR   PDB; 1KOO; 27-FEB-02.
DR   PDB; 1OAI; 20-FEB-03.
DR   Genew; HGNC:8071; NXF1.

  Query Match             7.6%;  Score 93;  DB 1;  Length 619;
  Best Local Similarity   24.7%;  Pred. No. 0.65;
  Matches   60;  Conservative   32;  Mismatches   93;  Indels   58;  Gaps   13;

Qy          5 AGDSWGML------ACLCTVLWHLPAVPALNRTGDP-GPGPSIQKTYDLT---RYLEHQL 54
              :||  |:|      ||    :  :|  || :   :      :::|  | |    | |:|
Db        400 SGDRQGLLDAYHDGACCSLSIPFIPQNPARSSLAEYFKDSRNVKKLKDPTLRFRLLKHTR 459

Qy         55 RSLAGTYLNYLGPPFNEPDFNP--PRLGAETLPRATVDL-----EVWRSLNDKLR-LTQN 106
               ::   :|| |  |  : | |      : |:|        :     ||      | ||  |:
Db        460 LNVV-AFLNEL--PKTQHDVNSFVVDISAQTSTLLCFSVNGVFKEVDGKSRDSLRAFTRT 516

Qy        107 Y---EAYSHLLCYLRG--LNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQP 161
              :    | :  || :      | |:: |::|: |                      :| |
Db        517 FIAVPASNSGLCIVNDELFVRNASSEEIQRAFA----------------------MPAP 553

Qy        162 LPGTEPTWTPGPAHSDFLQK------MDDFWLLKELQTWLW---RSAKDFNRLKKKMQPP 212
               | : |  |  |   : ||       |:  |  | ||   |   |||: |  || | : |
Db        554 TPSSSPVPTLSPEQQEMLQAFSTQSGMNLEWSQKCLQDNNWDYTRSAQAFTHLKAKGEIP 613

Qy        213 AAA 215
                |
Db        614 EVA 616


RESULT 7
CTF1_HUMAN
ID   CTF1_HUMAN     STANDARD;      PRT;   201 AA.
AC   Q16619;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cardiotrophin-1 (CT-1).
GN   CTF1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart;
RX   MEDLINE=96429882; PubMed=8833032;
RA   Pennica D., Swanson T.A., Shaw K.J., Kuang W.-J., Gray C.L.,
RA   Beatty B.G., Wood W.I.;
RT   "Human cardiotrophin-1: protein and gene structure, biological and
RT   binding activities, and chromosomal localization.";
RL   Cytokine 8:183-189(1996).
CC   -!- FUNCTION: Induces cardiac myocyte hypertrophy in vitro. Binds to
CC       and activates the ILST/gp130 receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Highly expressed in heart, skeletal muscle,
CC       prostate and ovary. Lower levels in lung, kidney, pancreas,
CC       thymus, testis and small intestine. Little or no expression in
CC       brain, placenta, liver, spleen, colon or peripheral blood
CC       leukocytes.
CC   -!- SIMILARITY: Belongs to the IL-6 superfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U43033; AAD12173.1; -.
DR   EMBL; U43031; AAD12173.1; JOINED.
DR   EMBL; U43032; AAD12173.1; JOINED.
DR   EMBL; U43030; AAA85229.1; -.
DR   PIR; G02312; G02312.
DR   Genew; HGNC:2499; CTF1.
DR   MIM; 600435; -.
DR   GO; GO:0005576; C:extracellular; TAS.
DR   GO; GO:0005146; F:leukemia inhibitory factor receptor binding; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0007517; P:muscle development; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
KW   Cytokine; Polymorphism.
FT   VARIANT      92      92       A -> T (in dbSNP:2234933).
FT                                 /FTId=VAR_014938.
SQ   SEQUENCE   201 AA;  21227 MW;  0235A7B5745F675F CRC64;

  Query Match             7.5%;  Score 91.5;  DB 1;  Length 201;
  Best Local Similarity   25.9%;  Pred. No. 0.21;
  Matches   45;  Conservative   24;  Mismatches   92;  Indels   13;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLND 99
              |::|: |   |       |    |:   | ||  | |:||||    |  |         :::
Db         27 IRQTHSLAHLLTKYAEQLLQEYVQLQGDPFGLPSFSPPRLPVAGL-SAPAPSHAGLPVHE 85

Qy        100 KLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVMAAL 154
              :||  :  | : |    |  : |: |         | | |         : |  ::  ::|||
Db         86 RLRL--DAAALAALPPLLDAVCRRQAELNPRAPRLLRRLEDAARQARALGAAVEALLAAL 143

Qy        155 GYPLPQPLPGTEP---TWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              |      |  ||   | :   |   |  |:    :    : || |:  |  :|
Db        144 G--AANRGPRAEPPAATASAASATGVFPAKVLGLRVCGLYREWLSRTEGDLGQL 195


RESULT 8
HAIR_MOUSE
ID   HAIR_MOUSE     STANDARD;      PRT;  1182 AA.
AC   Q61645; Q80Y47;
DT   01-NOV-1997 (Rel. 35, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hairless protein.
GN   HR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Skin;
RX   MEDLINE=94329587; PubMed=8052649;
RA   Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,
RA   Stoye J.P.;
RT   "Structure and expression of the hairless gene of mice.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Retina;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
```

```
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May act as a transcription factor that could act on to
CC       regulate one of the phases of hair growth.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Expressed predominantly in brain, hair
CC       follicles and interfollicular epidermis. No expression in
CC       dermis.
CC   -!- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS
CC       INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR
CC       FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH
CC       CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER
CC       MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED
CC       RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMOZYGOUS FOR
CC       HR MUTATION ARE UNIQUELY SENSITIVE TO UV AND CHEMICALLY INDUCED
CC       SKIN TUMORS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z32675; CAA83587.1; -.
DR   EMBL; BC049182; AAH49182.1; -.
DR   PIR; I48378; I48378.
DR   MGD; MGI:96223; hr.
DR   InterPro; IPR003347; TF_JmjC.
DR   Pfam; PF02373; jmjC; 1.
DR   SMART; SM00558; JmjC; 1.
KW   Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
KW   Metal-binding.
FT   ZN_FING      595    620       C6-TYPE.
FT   DOMAIN       535    540       POLY-GLY.
FT   CONFLICT     401    401       P -> S (IN REF. 1).
SQ   SEQUENCE   1182 AA;  127192 MW;  3AFABE96C6EB3241 CRC64;

  Query Match             7.5%;  Score 91.5;  DB 1;  Length 1182;
  Best Local Similarity   23.7%;  Pred. No. 2;
  Matches   70;  Conservative   25;  Mismatches   99;  Indels  101;  Gaps   17;

Qy          3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPG-------PGP---------------- 38
              ||||   |:  |      ||  |   |   ||       | |
Db        813 LRAGS--GLRKGL-----SLPLSPVRTRLSPPGALLWLQEPRPKHGFHLFQEHWRQGQPV 865

Qy         39 ---SIQKTYDLTRYLEHQLRSLAGTY--LNYLGPP------------FNEPDFNPPRLG 80
                |||   |: :     | :| |    |  ||||           |: |:  |
Db        866 LVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPPQPTNLDSTAFWEGFSHPETRP---- 921

Qy         81 AETLPRATVDLEVWRSLNDK-LRLTQN-----------YEAYSHLLCYL-RGLNRQAAT 126
                |   :| | : |:| ||      ||             ::   :|  ||  ||
Db        922 --KLDEGSV-LLLHRTLGDKDASRVQNLASSLPLPEYCAHQGKLNLASYLPLGLTLHPLE 978

Qy        127 AEL----------RRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAH 175
              :|           : |  :  |:  :    | |             | |    |
Db        979 PQLWAAYGVNSHRGHLGTKNLCVEVSDLISILVHAEAQL-----------PPWY--RAQ 1024

Qy        176 SDFLQKMD--DFWLLKELQTWLWR--SAKDFNRLKKKMQP--PAAAVTLHLGAHG 224
              |||  :|    |       : :|    |:|  |::: :|   || | ||  || |
Db       1025 KDFLSGLDGEGLWSPGSQTSTVWHVFRAQDAQRIRRFLQMVCPAGAGTLEPGAPG 1079


RESULT 9
CNTF_CHICK
ID   CNTF_CHICK     STANDARD;      PRT;   195 AA.
AC   Q02011;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Ciliary neurotrophic factor (CNTF) (Growth promoting activity) (GPA).
GN   CNF OR GPA.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92304573; PubMed=1610564;
RA   Leung D.W., Parent A.S., Cachianes G., Lee A.L., Nikolics K.,
RA   Esch F., Coulombe J.N., Blacher R.W., Eckenstein F.P., Nishi R.;
RT   "Cloning, expression during development, and evidence for release of
RT   a trophic factor for ciliary ganglion neurons.";
RL   Neuron 8:1045-1053(1992).
CC   -!- FUNCTION: CNTF is a survival factor for various neuronal cell
CC       types. Seems to prevent the degeneration of motor axons after
CC       axotomy.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Nervous system.
CC   -!- SIMILARITY: Belongs to the CNTF family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80827; AAA48784.1; -.
DR   PIR; JH0680; JH0680.
DR   HSSP; P26441; 1CNT.
DR   InterPro; IPR000151; Cil_neuro_factor.
DR   Pfam; PF01110; CNTF; 1.
DR   ProDom; PD011041; Cil_neuro_factor; 1.
KW   Growth factor; Neurone.
SQ   SEQUENCE   195 AA;  21330 MW;  FEA076949DB34AC5 CRC64;

  Query Match             7.4%;  Score 91;  DB 1;  Length 195;
  Best Local Similarity   27.3%;  Pred. No. 0.23;
  Matches   51;  Conservative   21;  Mismatches   83;  Indels   32;  Gaps    9;

Qy         46 LTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQ 105
              | | :   :  |   |:   |       | :       : :| | |  | |        ||
Db         23 LARKMRSDVTDLLDIYVERQG-----LDASISVAAVDGVPTAAV--ERWAEQTGTQRLLD 75

Qy        106 N---YEAYSHLLCYLRGLNRQA---ATAELRRSLA-------HFCTSLQGLLGSIAGVMA 152
              |   | |:  ||  :    |:      |||  :||         |   |: ||      :
Db         76 NLAAYRAFRTLLAQMLEEQRELLGDTDAELGPALAAMLLQVSAFVYHLEELL-----ELE 130

Qy        153 ALGYPLPQPLPGTEPTWTPGPAH-SDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQP 211
              : | |  :   |:|| | |  | | ||:    :|:||  |  || :|  :| | |
Db        131 SRGAPAEE---GSEP---PAPPRLSLFEQKLRGLRVLRELAQWAVRSVRDLRQLSKHGPG 184

Qy        212 PAAAVTL 218
                ||: |
Db        185 SGAALGL 191
```

```
RESULT 10
MIP1_SCHPO
ID   MIP1_SCHPO     STANDARD;      PRT;  1313 AA.
AC   P87141;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   WD-repeat protein mip1.
GN   MIP1 OR SPAC57A7.11.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RX   MEDLINE=20115869; PubMed=10648609;
RA   Shinozaki-Yabana S., Watanabe Y., Yamamoto M.;
RT   "Novel WD-repeat protein Mip1p facilitates function of the meiotic
RT   regulator Mei2p in fission yeast.";
RL   Mol. Cell. Biol. 20:1234-1242(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: BINDS TO AND FACILITATES THE FUNCTIONING OF THE MEIOTIC
CC       REGULATOR MEI2. MAY ALSO BE INVOLVED IN CONJUGATION BY INTERACTING
CC       WITH STE11. ESSENTIAL FOR CELL GROWTH.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Contains 7 WD repeats.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB032552; BAA84585.1; -.
DR   EMBL; Z95396; CAB08769.1; -.
DR   PIR; T38943; T38943.
DR   GeneDB_SPombe; SPAC57A7.11; -.
DR   InterPro; IPR008938; ARM.
DR   InterPro; IPR001680; WD40.
DR   InterPro; IPR004083; Yeast176.
DR   Pfam; PF00400; WD40; 5.
DR   PRINTS; PR01547; YEAST176DUF.
DR   SMART; SM00320; WD40; 6.
DR   PROSITE; PS00678; WD_REPEATS_1; 1.
DR   PROSITE; PS50082; WD_REPEATS_2; 2.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Meiosis; WD repeat; Repeat.
FT   REPEAT      986   1029       WD 1.
FT   REPEAT     1033   1074       WD 2.
FT   REPEAT     1087   1126       WD 3.
FT   REPEAT     1130   1170       WD 4.
FT   REPEAT     1176   1216       WD 5.
FT   REPEAT     1219   1259       WD 6.
FT   REPEAT     1268   1308       WD 7.
SQ   SEQUENCE   1313 AA;  148533 MW;  C71B663B0171E7A4 CRC64;

  Query Match             7.3%;  Score 89.5;  DB 1;  Length 1313;
  Best Local Similarity   22.8%;  Pred. No. 3.6;
  Matches   61;  Conservative   32;  Mismatches   69;  Indels  105;  Gaps   14;

Qy         15 LCTVLWHLPAV--PALNR-------TGDPGPGPSI---------QKTYDLTRYLEHQLRS 56
              |  :| ||||:   :|::       | || | | :           ::: ||| |:    |
Db        811 LAFLLQHLPALHKASLSKDTDTNSVTSDPKPHPFVPSVSENKILNRSFSLTRSLKGLALS 870

Qy         57 LAG-----------------TYLNYL------GPP-FNEPDFNPPRLGAETLPRATVDL 91
              |||                 : ||:|      ||| ||| ::           :: :|:
Db        871 LAGSDRASELLSLNGENKPAESNLNHLTSAKVPGPPAFNELEY----------QSELDM 919

Qy         92 EVWRSLND---------KLRLTQNYEAYSHLLCYLRGL---NRQAATAELRRSLAHFCTS 139
               :   | |          ::|  :: | |  | :|| : |   ||         | || : |:
Db        920 PLTSYLFDWSRKYFTEPQMRPNEDDEPGS--ICYNQRLWRRNRNEKLIYRTRPLAEYSTN 977

Qy        140 ------LQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKE--- 190
                   |    :||                          |    | | |    | :
Db        978 GRWNQQLMTFNNTIA----------------------PRKLMFHQFEDQLITLGDKDI 1013

Qy        191 LQTWLWRSAKDFNRLKKKMQPPAAAVT 217
              :| | ||     ||     :  |:| |
Db       1014 IQVWDWRR----NRCLNSFKTSASATT 1036


RESULT 11
HAIR_RAT
ID   HAIR_RAT       STANDARD;      PRT;  1181 AA.
AC   P97609;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Hairless protein.
GN   HR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=97141510; PubMed=8987811;
RA   Thompson C.C.;
RT   "Thyroid hormone-responsive genes in developing cerebellum include a
RT   novel synaptotagmin and a hairless homolog.";
RL   J. Neurosci. 16:7832-7840(1996).
CC   -!- FUNCTION: May act as a transcription factor that could act on to
CC       regulate one of the phases of hair growth.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; U71293; AAC53018.1; ALT_INIT.
DR   InterPro; IPR003347; TF_JmjC.
DR   Pfam; PF02373; jmjC; 1.
DR   SMART; SM00558; JmjC; 1.
KW   Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
KW   Metal-binding.
FT   ZN_FING     594    619       C6-TYPE.
SQ   SEQUENCE   1181 AA;  127307 MW;  834B7029CF8E88F0 CRC64;

  Query Match             7.1%;  Score 87.5;  DB 1;  Length 1181;
  Best Local Similarity   25.3%;  Pred. No. 4.8;
  Matches   61;  Conservative   21;  Mismatches   92;  Indels   67;  Gaps   12;

Qy         30 RTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTY--LNYLGPP----------FNEPDF 74
Db        859 RQGQPVLVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPPQPTSLDSTAFWKGFSHPEA 918

Qy         75 NPPRLGAETLPRATVDLEVWRSLNDK----------LRLTQ--------NYEAYSHLLC 115
Db        919 RP------KLDEGSV-LLLHRPLGDKDESRVENLASSLPLPEYCAHQGKLNLASYLPLGL 971

Qy        116 YLRGLNRQAATAELRRS------LAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTW 169
Db        972 TLHPLEPQLWAAYGVNSHRGHLGTKNLCVEVSDLISILVHAEAQL----------PPW 1019

Qy        170 TPGPAHSDFLQKMD--DFWLLKELQTWLWR--SAKDFNRLKKKMQP--PAAAVTLHLGAH 223
Db       1020 Y--RAQKDFLSGLDGEGLWSPGSQTSTVWHVFRAQDAQRIRRFLQMVCPAGAGTLEPGAP 1077

Qy        224 G 224
Db       1078 G 1078


RESULT 12
M2B2_HUMAN
ID   M2B2_HUMAN     STANDARD;      PRT;  1009 AA.
AC   Q9Y2E5;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Epididymis-specific alpha-mannosidase precursor (EC 3.2.1.24)
DE   (Mannosidase alpha class 2B member 2).
GN   MAN2B2 OR KIAA0935.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stone N.E., Schmutz J.J., Cox D.R., Myers R.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 132-1009 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=99246063; PubMed=10231032;
RA   Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA   Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XIII.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:63-70(1999).
RN   [3]
RP   RECONSTRUCTION FROM GENOMIC SEQUENCE.
RA   Bairoch A.;
RL   Unpublished observations (NOV-2001).
CC   -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing alpha-D-
CC       mannose residues in alpha-D-mannosides.
CC   -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC   -!- SIMILARITY: Belongs to family 38 of glycosyl hydrolases.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AC004480; -; NOT_ANNOTATED_CDS.
DR   EMBL; AB023152; BAA76779.1; ALT_SEQ.
DR   InterPro; IPR000602; Glyco_hydro_38.
DR   Pfam; PF01074; Glyco_hydro_38; 1.
KW   Hydrolase; Glycosidase; Signal; Glycoprotein.
FT   SIGNAL       1     23       POTENTIAL.
FT   CHAIN       24   1009       EPIDIDYMIS-SPECIFIC ALPHA-MANNOSIDASE.
FT   CARBOHYD   226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   249    249       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   516    516       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   608    608       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   670    670       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   675    675       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   748    748       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   808    808       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   812    812       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   890    890       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1009 AA;  113987 MW;  F3DB81DD061352E6 CRC64;

  Query Match             7.0%;  Score 86;  DB 1;  Length 1009;
  Best Local Similarity   23.3%;  Pred. No. 5.5;
  Matches   45;  Conservative   20;  Mismatches   66;  Indels   62;  Gaps    8;

Qy         50 LEHQ----LRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQ 105
Db        842 LQHRPVVLFGDLAGTAPKLPGPQQQEAVTLPPNLHLQIL-----SIPGWRYSSNHTEHSQ 896

Qy        106 NYEAYSHLLCYLRGLNRQAATAELRRSLAHF----------------CTSLQGLLGSIAG 149
Db        897 N----------LRKGHRGEAQADLRRVLLRLYHLYEVGEDPVLSQPVTVNLQAVLQALGS 946

Qy        150 VMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNR--LKK 207
Db        947 VVAV----EERSLTGT-------------------WDLSMLHRWSWRTGPGRHRGDTTS 982

Qy        208 KMQPPAA-AVTLH 219
Db        983 PSRPPGGPIITVH 995


RESULT 13
PHAC_PSEOL
ID   PHAC_PSEOL     STANDARD;      PRT;   560 AA.
AC   P26496;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Poly(3-hydroxyalkanoate) polymerase 2 (EC 2.3.1.-) (PHA polymerase 2)
DE   (PHA synthase 2) (Polyhydroxyalkanoic acid synthase 2).
GN   PHAC.
OS   Pseudomonas oleovorans.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=301;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GPo1;
RX   MEDLINE=91115830; PubMed=1989978;
RA   Huisman G.W., Wonink E., Meima R., Kazemier B., Terpstra P.,
RA   Witholt B.;
RT   "Metabolism of poly(3-hydroxyalkanoates) (PHAs) by Pseudomonas
RT   oleovorans. Identification and sequences of genes and function of the
RT   encoded proteins in the synthesis and degradation of PHA.";
RL   J. Biol. Chem. 266:2191-2198(1991).
CC   -!- FUNCTION: P.OLEOVORANS ACCUMULATES POLY(3-HYDROXYALKANOATES) AFTER
```

```
CC       GROWTH ON MEDIUM CHAIN LENGTH HYDROCARBONS. LARGE AMOUNTS OF THIS
CC       POLYESTER ARE SYNTHESIZED WHEN CELLS ARE GROWN UNDER NITROGEN-
CC       LIMITING CONDITIONS. WHEN NITROGEN IS RESUPPLIED IN THE MEDIUM,
CC       THE ACCUMULATED PHA IS DEGRADED.
CC   -!- SIMILARITY: Belongs to the PHA/PHB synthase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M58445; AAA25934.1; -.
DR   PIR; C38604; C38604.
DR   InterPro; IPR000073; A/b_hydrolase.
DR   Pfam; PF00561; abhydrolase; 1.
KW   PHA biosynthesis; Transferase; Acyltransferase.
FT   ACT_SITE    296    296       POTENTIAL.
SQ   SEQUENCE   560 AA;  62631 MW;  E2CD844FC1616B83 CRC64;

  Query Match             7.0%;  Score 85.5;  DB 1;  Length 560;
  Best Local Similarity   30.2%;  Pred. No. 2.9;
  Matches   38;  Conservative   17;  Mismatches   42;  Indels   29;  Gaps    7;

Qy          77 PRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRRSL--A 134
               |  |  |||  :::::     |  | |          |:  || ::||:    ||  |  |
Db           5 PAKGTPTLPATSMNVQ-----NAILGLRGR-----DLISTLRNVSRQS----LRHPLHTA 50

Qy         135 HFCTSLQGLLGSIAGVMAALGYPLPQPLPG----TEPTWTPGPAHSDFLQKMDDFWLLKE 190
               |   :| | || :      ||    || |     ::|||:  | :   ||     :|  :
Db          51 HHLLALGGQLGRV-----ILGDTPLQPNPRDPRFSDPTWSQNPFYRRGLQA----YLAWQ 101

Qy         191 LQTWLW 196
               || ||
Db         102 KQTRLW 107


RESULT 14
CNTF_PIG
ID   CNTF_PIG      STANDARD;      PRT;   200 AA.
AC   O02732;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Ciliary neurotrophic factor (CNTF).
GN   CNTF.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98030048; PubMed=9363597;
RA   Weaks R.L., Ramsoondar J.J., Gallagher D.S. Jr., Nogues C.,
RA   Piedrahita J.A.;
RT   "Isolation, characterization and chromosomal localization of the
RT   porcine ciliary neurotrophic factor (CNTF) gene.";
RL   Anim. Genet. 28:354-357(1997).
CC   -!- FUNCTION: CNTF is a survival factor for various neuronal cell
CC       types. Seems to prevent the degeneration of motor axons after
CC       axotomy (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: Nervous system (By similarity).
CC   -!- SIMILARITY: Belongs to the CNTF family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U57644; AAC27342.1; -.
DR   HSSP; P26441; 1CNT.
DR   InterPro; IPR000151; Cil_neuro_factor.
DR   Pfam; PF01110; CNTF; 1.
DR   ProDom; PD011041; Cil_neuro_factor; 1.
KW   Growth factor; Neurone.
SQ   SEQUENCE   200 AA;  22718 MW;  40507C4457ED6531 CRC64;

  Query Match             6.9%;  Score 84;  DB 1;  Length 200;
  Best Local Similarity   24.7%;  Pred. No. 1.1;
  Matches   47;  Conservative   23;  Mismatches   88;  Indels   32;  Gaps    9;

Qy          46 LTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQ 105
               | | :   | :|   |: : |   ||   |       : :| |: |   |  | :  || :
Db          23 LARKIRSDLTALMEAYVKHQG--LNE---NINLDSVDGVPMASTDR--WSELTEAERLQE 75

Qy         106 NYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTS---LQGLLGSIAGVMAALGYPLPQ-- 160
               |  ||      |  |        |    ||  :       : :|   :||  | | :
Db          76 NLRAYRTFHVMLARL--------LEDQREHFTPAEDDFHQAIHTIVLQVAAFAYQLEELM 127

Qy         161 -----PLPGTEPTWTPGPAHSD--FLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPA 213
                    :| :|   ||       | :|:    :|:||  |  || :|   :
Db         128 VLLEHKVPPSEADGTPLSVGGGGLFEKKLWGLKVLQELSQWTVRSIRDLRVISSHQ---- 183

Qy         214 AAVTLHLGAH 223
               | |  | |:|
Db         184 AGVPAH-GSH 192


RESULT 15
TLL_DROME
ID   TLL_DROME      STANDARD;      PRT;   452 AA.
AC   P18102; Q9VA33;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tailless protein.
GN   TLL OR NR2E2 OR CG1378.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Oregon-R;
RX   MEDLINE=90304905; PubMed=2364433;
RA   Pignoni F., Baldarelli R.M., Steingrimsson E., Diaz R.J.,
RA   Patapoutian A., Merriam J.R., Lengyel J.A.;
RT   "The Drosophila gene tailless is expressed at the embryonic termini
RT   and is a member of the steroid receptor superfamily.";
RL   Cell 62:151-163(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93157371; PubMed=8430097;
RA   Liaw G.-J., Steingrimsson E., Pignoni F., Courey A.J., Lengyel J.A.;
RT   "Characterization of downstream elements in a Raf-1 pathway.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:858-862(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
```

```
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=99287803; PubMed=10357938;
RA   Daniel A., Dumstrei K., Lengyel J.A., Hartenstein V.;
RT   "The control of cell fate in the embryonic visual system by atonal,
RT   tailless and EGFR signaling.";
RL   Development 126:2945-2954(1999).
CC   -!- FUNCTION: ORPHAN RECEPTOR THAT BINDS DNA AS A MONOMER TO HORMONE
CC       RESPONSE ELEMENTS (HRE) CONTAINING AN EXTENDED CORE MOTIF HALF-
CC       SITE SEQUENCE 5'-AAGTCA-3' IN WHICH THE 5' FLANKING NUCLEOTIDES
CC       PARTICIPATE IN DETERMINING RECEPTOR SPECIFICITY. THIS RECEPTOR
CC       BINDS TO THE CONSENSUS SEQUENCE [AG][AG]AAGTCAA. PLAYS A KEY ROLE
CC       IN THE ESTABLISHMENT OF NONMETAMERIC DOMAINS AT THE ANTERIOR AND
CC       POSTERIOR POLES OF THE EMBRYO. IT MAY ALSO PLAY A ROLE IN THE
CC       NERVOUS SYSTEM. THE MATERNAL TERMINAL PATHWAY ACTIVATES THE TLL
CC       GENE IN THE TERMINI; TLL ACTIVITY THEN REPRESSES SEGMENTATION AND
CC       ACTIVATES TERMINAL-SPECIFIC GENES IN THESE DOMAINS. INVOLVED IN
CC       THE REGULATION OF EARLY EYE DEVELOPMENT. IN THE EMBRYONIC VISUAL
CC       SYSTEM ANLAGE DRIVES CELLS TO OPTIC LOBE AS OPPOSED TO BOLWIG'S
CC       ORGAN FATE.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- TISSUE SPECIFICITY: BRAIN AND PERIPHERAL NERVOUS SYSTEM.
CC   -!- DEVELOPMENTAL STAGE: DURING STAGE 10 FOUND IN THE ANTERIOR PART OF
CC       THE VISUAL SYSTEM THAT LATER GIVES RISE TO THE ANTERIOR LIP OF THE
CC       OPTIC LOBE. AT STAGE 12 ALSO FOUND IN THE POSTERIOR LIP OF THE
CC       OPTIC LOBE. IN THIRD LARVAL INSTAR EXPRESSED IN THE OPTIC LOBE OF
CC       THE LARVAL BRAIN AND IN THE EYE ANTENNAL DISK, BOTH IN ANTENNAL
CC       AND EYE PORTION.
CC   -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR2
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M34639; AAA28936.1; -.
DR   EMBL; AF019362; AAB71371.1; -.
DR   EMBL; AE003775; AAF57091.1; -.
DR   PIR; A35602; A35602.
DR   HSSP; P10826; 1HRA.
DR   TRANSFAC; T00789; -.
DR   FlyBase; FBgn0003720; tll.
DR   GO; GO:0004879; F:ligand-dependent nuclear receptor activity; NAS.
DR   GO; GO:0007369; P:gastrulation; NAS.
DR   GO; GO:0007362; P:terminal region determination; IGI.
DR   GO; GO:0008293; P:torso signaling pathway; IGI.
DR   GO; GO:0006351; P:transcription, DNA-dependent; NAS.
DR   InterPro; IPR000536; Hormone_rec_lig.
DR   InterPro; IPR001723; Stdhrmn_receptor.
DR   InterPro; IPR008946; Str_ncl_receptor.
DR   InterPro; IPR001628; Znf_C4steroid.
DR   Pfam; PF00104; hormone_rec; 1.
DR   Pfam; PF00105; zf-C4; 1.
DR   PRINTS; PR00398; STRDHORMONER.
DR   PRINTS; PR00047; STROIDFINGER.
DR   ProDom; PD000035; Znf_C4steroid; 1.
DR   SMART; SM00430; HOLI; 1.
DR   SMART; SM00399; ZnF_C4; 1.
DR   PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW   Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW   Zinc-finger; Activator; Repressor; Developmental protein.
FT   DNA_BIND      34     101       NUCLEAR RECEPTOR-TYPE.
FT   ZN_FING       34      54       C4-TYPE.
FT   ZN_FING       70      96       C4-TYPE.
FT   DOMAIN       244     389       LIGAND-BINDING (BY SIMILARITY).
FT   DOMAIN       262     265       POLY-LEU.
SQ   SEQUENCE   452 AA;  50549 MW;  A4ABEFFDE993A37C CRC64;

  Query Match             6.8%;  Score 83;  DB 1;  Length 452;
  Best Local Similarity   23.8%;  Pred. No. 3.8;
  Matches   34;  Conservative   11;  Mismatches   32;  Indels   66;  Gaps    6;

Qy        115 CYLRGLNRQAATAE-------LRRSLAHFCTSLQGLLGSIAGVM-----------AALG 155
              |:  |:|: |   |       ||| :| :  :: |     || |           | |
Db         96 CFEVGMNKDAVQHERGPRNSTLRRHMAMYKDAMMG-----AGEMPQIPAEILMNTAALTG 150

Qy        156 YP-LPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAA 214
             :| :| |:||         |||    |                          ||:|
Db        151 FPGVPMPMPGLPQRAGHHPAHMAAFQ-------------------------PPPSA 181

Qy        215 AVTL-----------HLGAHGF 225
              |  |           | | |||
Db        182 AAVLDLSVPRVPHHPVHQGHHGF 204


Search completed: August 10, 2004, 06:46:26
Job time : 10 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:43:28 ; Search time 35 Seconds
(without alignments)
2028.331 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1226	100.0	225	4	Q9UBD9	Q9ubd9 homo sapien
2	1193	97.3	225	11	Q9QZM3	Q9qzm3 mus musculu
3	150.5	12.3	215	13	Q9PUJ1	Q9puj1 plethodon j
4	150.5	12.3	215	13	Q9PUJ0	Q9puj0 plethodon j
5	150.5	12.3	215	13	Q9PUJ2	Q9puj2 plethodon j
6	148.5	12.1	215	13	Q9PUI9	Q9pui9 plethodon j
7	105	8.6	471	16	Q88AR4	Q88ar4 pseudomonas
8	97.5	8.0	318	4	Q96LS2	Q96ls2 homo sapien
9	97.5	8.0	530	3	Q8X0E9	Q8x0e9 neurospora
10	97	7.9	8601	2	Q8GM87	Q8gm87 symbiont ba
11	96.5	7.9	243	4	Q8NEV9	Q8nev9 homo sapien
12	94	7.7	287	4	Q8N358	Q8n358 homo sapien
13	93	7.6	392	2	Q8KNF2	Q8knf2 micromonosp
14	92.5	7.5	860	16	Q82GI6	Q82gi6 streptomyce
15	92	7.5	332	10	Q9MAU1	Q9mau1 arabidopsis
16	91.5	7.5	716	16	Q8PI99	Q8pi99 xanthomonas
17	90	7.3	609	3	Q870A3	Q870a3 fusarium ox
18	89.5	7.3	771	2	Q9S3Q9	Q9s3q9 porphyromon
19	89	7.3	733	16	Q9I664	Q9i664 pseudomonas
20	88.5	7.2	298	10	Q84N16	Q84n16 zea mays (m
21	88	7.2	392	11	Q8BI13	Q8bi13 mus musculu
22	88	7.2	422	11	Q8BGR5	Q8bgr5 mus musculu
23	87	7.1	955	11	O88287	O88287 mus musculu
24	87	7.1	982	11	Q7TSJ4	Q7tsj4 mus musculu
25	87	7.1	1561	11	O88286	O88286 mus musculu
26	86.5	7.1	379	2	Q8KRX4	Q8krx4 streptomyce
27	86.5	7.1	640	16	O54153	O54153 streptomyce
28	85.5	7.0	522	7	Q8HWH4	Q8hwh4 mus musculu
29	85.5	7.0	1154	11	Q9Z1R2	Q9z1r2 mus musculu
30	85	6.9	476	11	Q8R363	Q8r363 mus musculu
31	85	6.9	727	11	O88841	O88841 mus musculu
32	84.5	6.9	294	11	Q9CPZ1	Q9cpz1 mus musculu
33	84.5	6.9	389	17	Q9YA71	Q9ya71 aeropyrum p
34	84	6.9	200	16	Q8YBB1	Q8ybb1 brucella me
35	84	6.9	232	16	Q8FX31	Q8fx31 brucella su
36	84	6.9	294	4	Q9NXX5	Q9nxx5 homo sapien
37	84	6.9	411	16	Q9RVF3	Q9rvf3 deinococcus
38	84	6.9	559	16	O83432	O83432 treponema p
39	84	6.9	655	11	Q8BY35	Q8by35 mus musculu
40	84	6.9	876	4	Q9P210	Q9p210 homo sapien
41	83.5	6.8	275	16	Q82E94	Q82e94 streptomyce
42	83.5	6.8	310	4	O95159	O95159 homo sapien
43	83.5	6.8	996	16	Q88GV1	Q88gv1 pseudomonas
44	83.5	6.8	5990	2	Q9RLP6	Q9rlp6 mycobacteri
45	83	6.8	222	2	Q9L584	Q9l584 streptococc

ALIGNMENTS

RESULT 1
Q9UBD9
ID Q9UBD9 PRELIMINARY; PRT; 225 AA.
AC Q9UBD9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Neurotrophin-1/B-cell stimulating factor-3 (Cardiotrophin-like
DE cytokine) (Similar to cardiotrophin-like cytokine, neurotrophin-1/B-
DE cell stimulating factor-3).
GN CLC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99432254; PubMed=10500198;
RA Senaldi G., Varnum B.C., Sarmiento U., Starnes C., Lile J., Scully S.,
RA Guo J., Elliott G., McNinch J., Shaklee C.L., Freeman D., Manu F.,
RA Simonet W.S., Boone T., Chang M.-S.;
RT "Novel neurotrophin-1/B cell-stimulating factor-3: A cytokine of the
RT IL-6 family.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99382254; PubMed=10448081;
RA Shi Y., Wang W., Yourey P.A., Gohari S., Zukauskas D., Zhang J.,
RA Ruben S., Alderson R.F.;
RT "Computational EST database analysis identifies a novel member of the
RT neuropoietic cytokine family.";
RL Biochem. Biophys. Res. Commun. 262:132-138(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Hu X., Xu Y., Zhang B., Peng X., Yuan J., Qiang B.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;

```
RA   Strausberg R.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF176912; AAF00992.1; -.
DR   EMBL; AF172854; AAD54284.1; -.
DR   EMBL; AF176911; AAF00991.1; -.
DR   EMBL; AY049779; AAL15436.1; -.
DR   EMBL; BC012939; AAH12939.1; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0005125; F:cytokine activity; NAS.
DR   GO; GO:0019221; P:cytokine and chemokine mediated signaling p. . .; NAS.
SQ   SEQUENCE   225 AA;  25176 MW;  E2DD4B6280833B55 CRC64;

  Query Match             100.0%;  Score 1226;  DB 4;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-107;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 2
Q9QZM3
ID   Q9QZM3      PRELIMINARY;      PRT;   225 AA.
AC   Q9QZM3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Neurotrophin-1/B-cell stimulating factor-3.
GN   BSF3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99432254; PubMed=10500198;
RA   Senaldi G., Varnum B.C., Sarmiento U., Starnes C., Lile J., Scully S.,
RA   Guo J., Elliott G., McNinch J., Shaklee C.L., Freeman D., Manu F.,
RA   Simonet W.S., Boone T., Chang M.-S.;
RT   "Novel neurotrophin-1/B cell-stimulating factor-3: a cytokine of the
RT   IL-6 family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:11458-11463(1999).
DR   EMBL; AF176913; AAF00993.1; -.
DR   MGD; MGI:1930088; Bsf3.
DR   GO; GO:0005125; F:cytokine activity; IDA.
DR   GO; GO:0030183; P:B-cell differentiation; IDA.
DR   GO; GO:0007166; P:cell surface receptor linked signal transdu. . .; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
SQ   SEQUENCE   225 AA;  25261 MW;  68B1FEAAB7F1A950 CRC64;

  Query Match             97.3%;  Score 1193;  DB 11;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 2e-104;
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||:|||||||:|||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||:|||||:||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||:|||||:||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 3
Q9PUJ1
ID   Q9PUJ1      PRELIMINARY;      PRT;   215 AA.
AC   Q9PUJ1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Receptivity factor isoform 2 precursor.
GN   PRF.
OS   Plethodon jordani (Jordan's salamander).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;
OC   Plethodon.
OX   NCBI_TaxID=8336;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99420364; PubMed=10489368;
RA   Rollmann S.M., Houck L.D., Feldhoff R.C.;
RT   "Proteinaceous pheromone affecting female receptivity in a terrestrial
RT   salamander.";
RL   Science 285:1907-1909(1999).
DR   EMBL; AF181481; AAF01026.1; -.
KW   Signal.
FT   SIGNAL       1      23       POTENTIAL.
SQ   SEQUENCE   215 AA;  24080 MW;  B341B8B7B4E28438 CRC64;

  Query Match             12.3%;  Score 150.5;  DB 13;  Length 215;
  Best Local Similarity   26.2%;  Pred. No. 4.1e-06;
  Matches   42;  Conservative   31;  Mismatches   80;  Indels    7;  Gaps    3;

Qy         56 SLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLC 115
              ||  |||:: | | ::||:  | :    || | :| : :    |: ||  |   || ::
Db         55 SLLPTYLSFQGAPLSDPDYQLPHIKVANLPTAAMDYDTFMKQTDETRLNNNLYFYSAIVE 114

Qy        116 YLR-GLNRQ----AATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWT 170
              :|:  :  |    |   |:         :   |:  |: :|  :|  :   || :|
Db        115 FLKEAMTEQEDLNPAELSLKAKFEEAMANSNTLISKISDIMTQMGMSVTITLP--KPLVV 172

Qy        171 PGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQ 210
              |    : | :|:    : || :  :  : :||  | || |
Db        173 PFEGSAYFRKKLRGGVVCKEYKERVLLTKRDFEFLAKKYQ 212


RESULT 4
Q9PUJ0
ID   Q9PUJ0      PRELIMINARY;      PRT;   215 AA.
AC   Q9PUJ0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Receptivity factor isoform 3 precursor.
GN   PRF.
OS   Plethodon jordani (Jordan's salamander).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;
OC   Plethodon.
OX   NCBI_TaxID=8336;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99420364; PubMed=10489368;
RA   Rollmann S.M., Houck L.D., Feldhoff R.C.;
RT   "Proteinaceous pheromone affecting female receptivity in a terrestrial
```

```
RT   salamander.";
RL   Science 285:1907-1909(1999).
DR   EMBL; AF181482; AAF01027.1; -.
KW   Signal.
FT   SIGNAL       1     23       POTENTIAL.
SQ   SEQUENCE   215 AA;  24024 MW;  6EB6E64AD93A1343 CRC64;

  Query Match             12.3%;  Score 150.5;  DB 13;  Length 215;
  Best Local Similarity   27.1%;  Pred. No. 4.1e-06;
  Matches   45;  Conservative   34;  Mismatches   72;  Indels   15;  Gaps    4;

Qy         56 SLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLC 115
              ||  |||:: | | ::||:  | :    || | :| : :    |: ||  |   || ::
Db         55 SLLPTYLSFQGAPLSDPDYQLPHIKVVNLPTAAMDYDTFIRQTDETRLKNNLYFYSAIVE 114

Qy        116 YLR-------GLN--RQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTE 166
              :|:       ||    |  |:  ::|:  |    |:  |: :|  :|  :   ||  :
Db        115 FLKEAMTEQEDLNPAELALKAKFEEAMANSNT----LISKISDIMTQMGMSVTITLP--K 168

Qy        167 PTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPP 212
              |   |    : | :|:    : || :  :: : :||  | :| | |
Db        169 PLVVPFKGSAYFSKKLRGGVVCKEYKERVFLTKRDFMLLAEKYQGP 214


RESULT 5
Q9PUJ2
ID   Q9PUJ2      PRELIMINARY;      PRT;   215 AA.
AC   Q9PUJ2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Receptivity factor isoform 1 precursor.
GN   PRF.
OS   Plethodon jordani (Jordan's salamander).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;
OC   Plethodon.
OX   NCBI_TaxID=8336;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99420364; PubMed=10489368;
RA   Rollmann S.M., Houck L.D., Feldhoff R.C.;
RT   "Proteinaceous pheromone affecting female receptivity in a terrestrial
RT   salamander.";
RL   Science 285:1907-1909(1999).
DR   EMBL; AF181480; AAF01025.1; -.
KW   Signal.
FT   SIGNAL       1     23       POTENTIAL.
SQ   SEQUENCE   215 AA;  24138 MW;  B1906BB666335738 CRC64;

  Query Match             12.3%;  Score 150.5;  DB 13;  Length 215;
  Best Local Similarity   26.2%;  Pred. No. 4.1e-06;
  Matches   42;  Conservative   31;  Mismatches   80;  Indels    7;  Gaps    3;

Qy         56 SLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLC 115
              ||  |||:: | | ::||:  | :    || | :| : :    |: ||  |   || ::
Db         55 SLLPTYLSFQGAPLSDPDYQLPHIKVANLPTAAMDYDTFMKQTDETRLNNNLYFYSAIVE 114

Qy        116 YLR-GLNRQ----AATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWT 170
              :|:  :  |      |   |:        :   |:  |: :|  :|  :   ||  :|
Db        115 FLKEAMTEQEDLNPAELSLKAKFEEAMANSNTLISKISDIMTQMGMSVTITLP--KPLVV 172

Qy        171 PGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQ 210
              |    : | :|:    : || :  :  : :||  | ||
Db        173 PFEGSAYFRKKLRGGVVCKEYKERVLLTKRDFEFLAKKYQ 212


RESULT 6
Q9PUI9
ID   Q9PUI9      PRELIMINARY;      PRT;   215 AA.
AC   Q9PUI9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Receptivity factor isoform 4 precursor.
GN   PRF.
OS   Plethodon jordani (Jordan's salamander).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Plethodontidae;
OC   Plethodon.
OX   NCBI_TaxID=8336;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99420364; PubMed=10489368;
RA   Rollmann S.M., Houck L.D., Feldhoff R.C.;
RT   "Proteinaceous pheromone affecting female receptivity in a terrestrial
RT   salamander.";
RL   Science 285:1907-1909(1999).
DR   EMBL; AF181483; AAF01028.1; -.
KW   Signal.
FT   SIGNAL       1     23       POTENTIAL.
SQ   SEQUENCE   215 AA;  24054 MW;  A4A412135FFA4E7F CRC64;

  Query Match             12.1%;  Score 148.5;  DB 13;  Length 215;
  Best Local Similarity   25.9%;  Pred. No. 6.3e-06;
  Matches   42;  Conservative   33;  Mismatches   80;  Indels    7;  Gaps    3;

Qy         56 SLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLC 115
              ||  |||:: | | ::||:  | :    || | :| : :    |: ||  |   || ::
Db         55 SLLPTYLSFQGAPLSDPDYRLPHIKVVNLPTAAMDYDTFIRQTDETRLKNNLYFYSAIVE 114

Qy        116 YLR-GLNRQ----AATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWT 170
              :|:  :  |      |   |:        :   |:  |: :|  :|  :   ||  :|
Db        115 FLKEAMTEQEDLNPAELSLKAKFEEAMANSNTLISKISDIMTQMGMSVTITLP--KPLVV 172

Qy        171 PGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPP 212
              |    : | :|:    : || :  :: : :||  | :| | |
Db        173 PFKGSAYFSKKLRGGVVCKEYKERVFLTKRDFMLLAEKYQGP 214


RESULT 7
Q88AR4
ID   Q88AR4      PRELIMINARY;      PRT;   471 AA.
AC   Q88AR4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Adenylate cyclase.
GN   PSPTO0322.
OS   Pseudomonas syringae (pv. tomato).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=323;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DC3000;
RA   Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA   Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA   Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA   Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA   White O., Fraser C., Collmer A.;
RT   "Complete sequence of Pseudomonas syringae.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE016857; AAO53867.1; -.
DR   TIGR; PSPTO0322; -.
DR   InterPro; IPR008172; Adenylate_cyc.
DR   Pfam; PF01928; CYTH; 1.
KW   Complete proteome.
SQ   SEQUENCE   471 AA;  53575 MW;  EF85D7CEC284D3AF CRC64;

  Query Match              8.6%;  Score 105;  DB 16;  Length 471;
  Best Local Similarity   25.0%;  Pred. No. 0.22;
  Matches   53;  Conservative   15;  Mismatches   98;  Indels   46;  Gaps    6;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ::|| ||   :|     :   || :|             | |     | |:    ||:
```

```
Db        182 LELREGDPAALLELAAELAATLPLMPC-----------DISKAERGYRLLDASSYSLS-- 228

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSH-------- 112

Db        229 -------------LPAPTLNAETPLDDAYSALAWHLLGSSQRLAEQYRFNGHWRLLQDWV 275

Qy        113 -LLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTP 171

Db        276 EMLGELRALTSSLGQAAPRTSTAQLRTALDALLED--------WRPLVQ--AGQEDADVR 325

Qy        172 GPAHSDFLQKMDDF-WLLKELQTWLWRSAKDF 202

Db        326 GAAHEQFLEELQDTRWGEFSLNTSRWLLARSW 357


RESULT 8
Q96LS2
ID   Q96LS2      PRELIMINARY;      PRT;   318 AA.
AC   Q96LS2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein FLJ25132.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA   Fukuzumi Y., Fujimori Y., Komiyama M., Suzuki Y., Hata H.,
RA   Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA   Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA   Nagai K., Isogai T., Sugano S.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK057861; BAB71598.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   318 AA;  34305 MW;  62E5CFB1A599BF40 CRC64;

  Query Match             8.0%;  Score 97.5;  DB 4;  Length 318;
  Best Local Similarity   25.1%;  Pred. No. 0.68;
  Matches   49;  Conservative   19;  Mismatches   76;  Indels   51;  Gaps    9;

Qy          6 GDSWGMLACLCTVLWHLPAVPALNRTG---DPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 62

Db         82 GTECSFLIYLSEVEWFCPPLPWRNQTAAQRAPKPPPKVQAVF----------RSNLSHLL 131

Qy         63 NYLGPPFNEPDFNPPRLGAETL-------PRATVDLEVW----RSLNDKLRLTQNYEAYS 111

Db        132 DLMGS------------GKESLIFMKKRTKRLTAQ---WALAAQRLAQKLGATQRDQ--K 174

Qy        112 HLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMA-----ALG-----YPLPQP 161

Db        175 QILVHIGFLTEESGDVFSPRVLMLFLCSALAWASALAPALAATLTLALGRALCRIPAPNP 234

Qy        162 LPGTEPTWTPGPAHS 176

Db        235 LPGSRSTPSPTPVES 249


RESULT 9
Q8X0E9
ID   Q8X0E9      PRELIMINARY;      PRT;   530 AA.
AC   Q8X0E9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   B14A6.080.
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   German Neurospora genome project;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL670007; CAD21310.1; -.
DR   InterPro; IPR001810; F-box.
DR   PROSITE; PS50181; FBOX; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   530 AA;  59352 MW;  70382EE15F71BB9D CRC64;

  Query Match             8.0%;  Score 97.5;  DB 3;  Length 530;
  Best Local Similarity   24.8%;  Pred. No. 1.3;
  Matches   41;  Conservative   22;  Mismatches   65;  Indels   37;  Gaps    6;

Qy         64 YLGPPFNEPDFNPPRLGAETLPRA-TVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 122

Db        236 FLVPPYLEPDFLPVLAGLKKLHLVLDADLATHRDMEIQYLLFKKFLALTPNLTWLR---- 291

Qy        123 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGT---------EPTWTPGP 173

Db        292 ----VNFERSISHY-----------GGKEALLKWLASSLRPGTWSSPGLINADPSRLPPP 336

Qy        174 AHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL 218

Db        337 VEFDHLEQLD----IGQLDV----SANTLYRLFNKFSSTLKAISL 373


RESULT 10
Q8GM87
ID   Q8GM87      PRELIMINARY;      PRT;  8601 AA.
AC   Q8GM87;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mixed type I polyketide synthase/nonribosomal peptide synthetase.
GN   PEDF.
OS   symbiont bacterium of Paederus fuscipes.
OC   Bacteria.
OX   NCBI_TaxID=176282;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22294974; PubMed=12381784;
RA   Piel J.;
RT   "A polyketide synthase-peptide synthetase gene cluster from an
RT   uncultured bacterial symbiont of Paederus beetles.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14002-14007(2002).
DR   EMBL; AY059471; AAL27851.1; -.
DR   GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
DR   GO; GO:0004459; F:L-lactate dehydrogenase activity; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR   GO; GO:0006096; P:glycolysis; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR   InterPro; IPR002198; ADH_short.
DR   InterPro; IPR002328; ADH_zinc.
DR   InterPro; IPR000873; AMP-bind.
DR   InterPro; IPR001242; Condensatn.
DR   InterPro; IPR002114; HPr_SerP_S.
DR   InterPro; IPR000794; Ketoacyl_synth.
DR   InterPro; IPR001557; L_LDH.
DR   InterPro; IPR001601; Methyltransf.
DR   InterPro; IPR006162; Ppantne_S.
```

```
DR   InterPro; IPR006163; Pp_bind.
DR   InterPro; IPR000051; SAM_bind.
DR   InterPro; IPR002155; Thiolase.
DR   Pfam; PF00106; adh_short; 1.
DR   Pfam; PF00501; AMP-binding; 1.
DR   Pfam; PF00668; Condensation; 1.
DR   Pfam; PF00109; ketoacyl-synt; 6.
DR   Pfam; PF02801; ketoacyl-synt_C; 6.
DR   Pfam; PF00550; pp-binding; 6.
DR   PROSITE; PS50075; ACP_DOMAIN; 6.
DR   PROSITE; PS00059; ADH_ZINC; 1.
DR   PROSITE; PS00455; AMP_BINDING; 1.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; 3.
DR   PROSITE; PS00064; L_LDH; 1.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR   PROSITE; PS00589; PTS_HPR_SER; 1.
DR   PROSITE; PS00098; THIOLASE_1; 1.
SQ   SEQUENCE   8601 AA;  948046 MW;  C498AC8B0B6082C7 CRC64;

  Query Match             7.9%;  Score 97;  DB 2;  Length 8601;
  Best Local Similarity   25.6%;  Pred. No. 55;
  Matches   51;  Conservative   37;  Mismatches   69;  Indels   42;  Gaps   12;

Qy         58 AGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYL 117
              :|| :: :   |  :   ||: :  | :| : | |  ||:  | ::|  ::
Db       5078 SGTNVHLVVEEFVRSNSEDPRVLDDVSSTAQPEL-ILLSTKDAERLS---EVLNNLAHFV 5133

Qy        118 RGLNRQAATAELRRSLAHFCTSLQG-----------LLGSIAGVMAALGYPLPQPLPGTE 166
              |    | |  | | |||    :||           |:| :||:: ||   | :  |
Db       5134 RQAQNQPADLE-RLSLADLAYTLQTGREAMEQRVALLVGDLAGLLEALS-ALREERPCPV 5191

Qy        167 PTWT----PGPAH-----------SDFLQKMDDFWL----LKEL-QTWLWRSAKDFNRLK 206
                :    |||:           :: ||::   ||    | || | |:  :  |: :|:
Db       5192 SVWSGRVEPGPSRGAETVNADQPAAELLQRIPQ-WLAEGALDELAQAWVAGAPIDWCQLR 5250

Qy        207 KKMQPPAAAVTLHLGAHGF 225
              :: :||     :|| :: |
Db       5251 RR-RPPR---RVHLPSYPF 5265


RESULT 11
Q8NEV9
ID   Q8NEV9      PRELIMINARY;      PRT;   243 AA.
AC   Q8NEV9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   IL-27 p28 subunit.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pflanz S., Timans J., Cheung J., Rosales R., Kanzler H., Gilbert J.,
RA   Hibbert L., Churakova T., Travis M., Vaisberg E., Blumenschein W.,
RA   Mattson J., Wagner J., To W., Zurawski S., McClanahan T., Gorman D.,
RA   Bazan F., de Waal Malefyt R., Rennick D., Kastelein R.;
RT   "IL-27 p28 subunit sequences.";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY099296; AAM34498.1; -.
DR   Genew; HGNC:19157; IL30.
DR   GO; GO:0005615; C:extracellular space; ISS.
DR   GO; GO:0045523; F:interleukin-27 receptor binding; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0005102; F:receptor binding; ISS.
DR   GO; GO:0045078; P:positive regulation of interferon-gamma bio. . .; IDA.
DR   GO; GO:0042129; P:regulation of T-cell proliferation; ISS.
DR   GO; GO:0045625; P:regulation of T-helper 1 cell differentiation; IDA.
SQ   SEQUENCE   243 AA;  27476 MW;  2FA8A8EB6CF61AE8 CRC64;

  Query Match             7.9%;  Score 96.5;  DB 4;  Length 243;
  Best Local Similarity   23.7%;  Pred. No. 0.6;
  Matches   58;  Conservative   26;  Mismatches  104;  Indels   57;  Gaps    9;

Qy          1 MDLRAGD-SWGMLACLCTVL------WHLPAVPALNRTGDPGPGPSIQK-------TYDL 46
              |   |||  | :   | :|      | | | |          |  |:|:       : |
Db          1 MGQTAGDLGWRLSLLLLPLLLVQAGVWGFPRPPG-------RPQLSLQELRREFTVSLHL 53

Qy         47 TRYLEHQLRSLAGTYLN--------YLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLN 98
               | |  ::|  |  :           || |           || | ||  ::  : || |:
Db         54 ARKLLSEVRGQAHRFAESHLPGVNLYLLP-----------LG-EQLPDVSLTFQAWRRLS 101

Qy         99 DKLRLTQNYEAYSHLLCYLRGLNRQAATAELRR-SLAHFCTSLQGLLGSIAGVMAALGYP 157
              |  ||                | ||  |     : |  |       |: |   :   : | |:
Db        102 DPERLCFISTTLQPFHAPLGGLGTQGRWTNMERMQLWAMRLDLRDLQRHLRFQVLAAGFN 161

Qy        158 LPQP--------------LPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDF 202
              ||:               |||   :   |||   : | :  : ||  |:  | |: ::
Db        162 LPEEEEEEEEEEEEERKGLLPGALGSALQGPAQVSWPQLLSTYRLLHSLELVLSRAVREL 221

Qy        203 NRLKK 207
              | |
Db        222 LLLSK 226


RESULT 12
Q8N358
ID   Q8N358      PRELIMINARY;      PRT;   287 AA.
AC   Q8N358;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Blood;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC028076; AAH28076.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   287 AA;  32156 MW;  6E571ECA0328E74E CRC64;

  Query Match             7.7%;  Score 94;  DB 4;  Length 287;
  Best Local Similarity   22.3%;  Pred. No. 1.3;
  Matches   47;  Conservative   18;  Mismatches   62;  Indels   84;  Gaps    9;

Qy         15 LCTVLWHLPAVPALNRTGDPGPGPSIQKTYD----------LTRY---LEHQL--RSLAG 59
              |  ::|  |          | |||::  |           : |:   | | |  |:
Db        132 LLPLVWLRP----------PDPGPTLNLTLQCPGSRQSPPGVLRFHNQLWHVLTGRAFTN 181

Qy         60 TYLNY----LGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLC 115
              ||| |    :||                                ::   ||  |
Db        182 TYLFYGAYRVGP-----------------------------------ESSSVYSIRLA 204

Qy        116 YLRGLNRQAATAELRRSLAHFCTSLQGLLGSIA------GVMAALGYPLPQPLPGTEPTW 169
              ||         : |   |  || :|  || ::|      |  |   :|   ||||
Db        205 YL--------LSPLACLLLCFCGTLHQLLSTLASGPGAGGPWAPAHWPACPPLPGLPRLQ 256

Qy        170 TPGPAHSDFLQKM----DDFWLLKELQTWLW 196
              | | |:  |:       | |  :    | |
Db        257 LPPPHHAQVLRAAGAMGGDTWRGRSFLPWGW 287


RESULT 13
Q8KNF2
ID   Q8KNF2      PRELIMINARY;      PRT;   392 AA.
AC   Q8KNF2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CalG1.
```

```
GN   CALG1.
OS   Micromonospora echinospora (Micromonospora purpurea).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Micromonosporineae; Micromonosporaceae; Micromonospora.
OX   NCBI_TaxID=1877;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL 15839;
RA   Ahlert J., Shepard E., Lomovskaya N., Zazopoulos E., Staffa A.,
RA   Bachmann B.O., Huang K., Fonstein L., Czisny A., Whitwam R.E.,
RA   Farnet C.M., Thorson J.S.;
RT   "The calicheamicin gene cluster and its iterative type I PKS.";
RL   Science 0:0-0(2002).
DR   EMBL; AF497482; AAM70336.1; -.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0016301; F:kinase activity; IEA.
DR   GO; GO:0016774; F:phosphotransferase activity, carboxyl group. . .; IEA.
DR   GO; GO:0016758; F:transferase activity, transferring hexosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   GO; GO:0030259; P:lipid glycosylation; IEA.
DR   GO; GO:0016310; P:phosphorylation; IEA.
DR   InterPro; IPR000890; Acetate_kin.
DR   InterPro; IPR004276; Glyco_trans_28.
DR   Pfam; PF03033; Glyco_transf_28; 1.
DR   PROSITE; PS01076; ACETATE_KINASE_2; 1.
SQ   SEQUENCE   392 AA;  41151 MW;  8D298611281E065E CRC64;

  Query Match             7.6%;  Score 93;  DB 2;  Length 392;
  Best Local Similarity   26.4%;  Pred. No. 2.4;
  Matches   43;  Conservative   15;  Mismatches   61;  Indels   44;  Gaps    7;

Qy          45 DLTRYLEHQLRSLA----------------GTYLNYLGPPFNEPDF--NPPRLGAETLPR 86
               |||| :| ::| ||                  :::   |   ||:|   | |     :|
Db         146 DLTRSIEEEVRGLAQRLGLDLPPGRIDGFGNPFIDIFPPSLQEPEFRARPRRHELRPVPF 205

Qy          87 ATV-DLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLG 145
               |   ||  | | | | |                | ||       | |: |:      :: || |
Db         206 AEQGDLPAWLSSRDTAR----------PLVYLTLGTSSGGTVEVLRA------AIDGLAG 249

Qy         146 SIAGVMAALGYPLPQPLPGTEP------TWTPGPA---HSDFL 179
                 | |: | |  |    |  |      :| |  |   | | :
Db         250 LDADVLVASGPSLDVSGLGEVPANVRLESWVPQAALLPHVDLV 292


RESULT 14
Q82GI6
ID   Q82GI6      PRELIMINARY;      PRT;   860 AA.
AC   Q82GI6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   SAV3911.
OS   Streptomyces avermitilis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=33903;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=21477403; PubMed=11572948;
RA   Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA   Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA   Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT   "Genome sequence of an industrial microorganism Streptomyces
RT   avermitilis: deducing the ability of producing secondary
RT   metabolites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX   MEDLINE=22608306; PubMed=12692562;
RA   Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA   Sakaki Y., Hattori M., Omura S.;
RT   "Complete genome sequence and comparative analysis of the industrial
RT   microorganism Streptomyces avermitilis.";
RL   Nat. Biotechnol. 21:526-531(2003).
DR   EMBL; AP005037; BAC71623.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   860 AA;  90116 MW;  C7F0123F70226BC5 CRC64;

  Query Match             7.5%;  Score 92.5;  DB 16;  Length 860;
  Best Local Similarity   27.2%;  Pred. No. 7.3;
  Matches   56;  Conservative   22;  Mismatches  101;  Indels   27;  Gaps    9;

Qy           5 AGDSWGMLACLCTVL--WHLPAVPALNRTGDPG------------PGPSIQKTYDLTRY 49
               || ::  || :  :| ||    ||: | |                | |:    :|| |
Db         288 AGQAYTALATVEELLKDWH-EGGPAVLRAGGLSVRDLKRTAVALDVPEPTAAFWVELT-Y 345

Qy          50 LEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEA 109
                 | |       |   |  :     |   ||   |       | :      |    :|
Db         346 AAGLLASDGEAEERYAATPAYDEWLELP--AAERWSRLAT---AWLAATRTAGLVGGRDA 400

Qy         110 YSHLLCYL-RGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPT 168
                 |  | ||:| ||:   || ||      :|   : | |:| | :   :|| || |:
Db         401 KERTLSALGPGLDRSAASEVRRRVLALLADLPEGASAAPASVLARLRW--ERPLRGTHPS 458

Qy         169 WTP--GPAHSDFLQKMDDFWLLKELQ 192
                 :|   |: || |:     | | | :
Db         459 PSPSSAPSSSDDLRTRLAEWTLSEAE 484


RESULT 15
Q9MAU1
ID   Q9MAU1      PRELIMINARY;      PRT;   332 AA.
AC   Q9MAU1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   F13M7.8 protein (Hypothetical protein).
GN   F13M7.8.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. columbia;
RA   Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Kwan A., Liu, S,
RA   Li J., Kremenetskaia I., Luros J., Araujo R., Au M., Bredel V.,
RA   Buehler E., Conway A., Dewar K., Feng J., Kim C., Kurtz D., Li Y.,
RA   Palm C., Shinn P., Sun H., Davis R., Ecker J., Federspiel N.,
RA   Theologis A.;
RT   "The sequence of BAC F13M7 from Arabidopsis thaliana chromosome 1.";
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. columbia;
RA   Theologis A.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-Length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC004809; AAF40444.1; -.
DR   EMBL; AY085490; AAM62716.1; -.
```

```
DR   PIR; G86182; G86182.
KW   Hypothetical protein.
SQ   SEQUENCE   332 AA;  36793 MW;  18E8687141A070F4 CRC64;

  Query Match             7.5%;  Score 92;  DB 10;  Length 332;
  Best Local Similarity   26.4%;  Pred. No. 2.4;
  Matches   58;  Conservative   24;  Mismatches   72;  Indels   66;  Gaps   13;

Qy          23 PAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNY------------------ 64
               |:| | | :| | | ||   |||   | : |: ||   ::
Db         113 PSVTAGNLSGYP-PRPSF--TYDPGPYEQRQMESLLQQFIRERNPQIRPLPRLGLGSPVG 169

Qy          65 LGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQA 124
               |||    | |  ||:     | : :|     | | | |       :    |  :||  |:
Db         170 LGPIRASPQFLQPRVAPP--PTSILD----TSRNRKAR------SKDGALAVVRG--RKV 215

Qy         125 ATAELRRSL-----------AHFCTSLQGLLGSIAGVMAALGYPLPQPLP---GTEPTWT 170
                 |   ||           ||      |:    :|:|     |||:|||    || :
Db         216 RITEGSSSLYSLGRSWLKNGAHV-----GIQPQRSGIMK----PLPKPLPVDLTTETSVP 266

Qy         171 PGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQ 210
               |      :  :|   :|:|      | ||   ||: ::
Db         267 DDPDEESADEDKEDEEAVKQL------SEKDL--LKRHIE 298


Search completed: August 10, 2004, 06:47:48
Job time : 38 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:43:28 ; Search time 48 Seconds
(without alignments)
1324.441 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1226	100.0	225	2	AAW56141	Aaw56141 Amino aci
2	1226	100.0	225	2	AAW29715	Aaw29715 Human neu
3	1226	100.0	225	2	AAW94466	Aaw94466 Human car
4	1226	100.0	225	3	AAY87813	Aay87813 Human NNT
5	1226	100.0	225	4	AAG63543	Aag63543 Amino aci
6	1226	100.0	225	5	AAU78176	Aau78176 Human nov
7	1226	100.0	253	4	AAM25831	Aam25831 Human pro
8	1214	99.0	223	4	AAE00828	Aae00828 Human car
9	1204	98.2	321	4	ABB11896	Abb11896 Human car
10	1204	98.2	321	4	AAM79399	Aam79399 Human pro
11	1197	97.6	260	4	AAM78415	Aam78415 Human pro
12	1193	97.3	225	2	AAW56142	Aaw56142 Amino aci
13	1193	97.3	225	2	AAW29716	Aaw29716 Mouse neu
14	1193	97.3	225	3	AAY87814	Aay87814 Murine NN
15	1193	97.3	225	5	AAU78177	Aau78177 Mouse nov
16	1169	95.4	215	3	AAB19586	Aab19586 Human int
17	1136	92.7	215	3	AAB19587	Aab19587 Mouse int
18	885	72.2	164	4	AAM20115	Aam20115 Peptide #
19	885	72.2	164	4	ABB40317	Abb40317 Peptide #
20	885	72.2	164	4	AAM34012	Aam34012 Peptide #
21	885	72.2	164	4	ABB24716	Abb24716 Protein #
22	885	72.2	164	4	AAM73827	Aam73827 Human bon
23	885	72.2	164	4	AAM61118	Aam61118 Human bra
24	885	72.2	164	4	ABG55577	Abg55577 Human liv
25	885	72.2	164	5	ABG43716	Abg43716 Human pep
26	162.5	13.3	208	2	AAY09197	Aay09197 Human DNA
27	160	13.1	208	2	AAY09196	Aay09196 Human DNA
28	118.5	9.7	203	2	AAR83965	Aar83965 Mouse car
29	118.5	9.7	203	2	AAR88204	Aar88204 Human car
30	118.5	9.7	203	2	AAW29237	Aaw29237 Murine ca
31	118.5	9.7	203	7	ADC24999	Adc24999 Mouse CHF
32	105	8.6	471	6	ABU41629	Abu41629 Protein e
33	97	7.9	8601	6	AAO29660	Aao29660 Paederus
34	96.5	7.9	243	4	AAU09153	Aau09153 Human cyt
35	96.5	7.9	243	4	AAB20277	Aab20277 Human int
36	96.5	7.9	243	5	ABG32291	Abg32291 Human IL-
37	96.5	7.9	243	5	AAU76375	Aau76375 Human hel
38	96.5	7.9	243	6	ABU09749	Abu09749 Human var
39	96.5	7.9	243	6	ABG75877	Abg75877 Human cyt
40	93	7.6	419	6	ABU12010	Abu12010 M. echino
41	92	7.5	332	3	AAG22132	Aag22132 Arabidops
42	92	7.5	332	3	AAG40321	Aag40321 Arabidops
43	91.5	7.5	201	2	AAR83967	Aar83967 Human car
44	91.5	7.5	201	2	AAW29238	Aaw29238 Human car
45	91.5	7.5	201	2	AAY06490	Aay06490 Human tum

ALIGNMENTS

RESULT 1
AAW56141
ID AAW56141 standard; protein; 225 AA.
XX
AC AAW56141;
XX
DT 13-JUL-1998 (first entry)
XX
DE Amino acid sequence of human neurotrophic factor NNT-1.
XX
KW Human; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW treatment; neurological disease; degeneration; Parkinson's disease;
KW amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1. .27
FT /note= "signal peptide"
FT Protein 28. .225
FT /note= "mature protein"
XX
PN US5741772-A.
XX
PD 21-APR-1998.
XX
PF 03-FEB-1997; 97US-00792019.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 1998-260526/23.
DR N-PSDB; AAV22652.
XX
PT Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT for stimulating growth of motor and sympathetic neurons.
XX
PS Claim 1; Fig 3; 41pp; English.
XX
CC The present sequence represents a human neurotrophic factor, designated
CC NNT-1, which is capable of stimulating growth of motor or sympathetic
CC neurons. The NNT-1 protein is useful in the treatment of neurological
CC diseases characterised by the degeneration and death of particular
CC classes of neurons. These diseases specifically include Parkinson's
CC disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke

```
CC   and various degenerative disorders affecting vision
XX
SQ   Sequence 225 AA;

  Query Match             100.0%;  Score 1226;  DB 2;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-119;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 2
AAW29715
ID   AAW29715 standard; protein; 225 AA.
XX
AC   AAW29715;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Human neurotrophic factor NNT-1.
XX
KW   NNT-1; neurotrophic factor; human; antiinflammatory; adjuvant;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   peripheral neurophathy; dystrophy; neural retina degeneration;
KW   common variable immunodeficiency; CVID; selective IgA deficiency;
KW   hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /label= Sig_peptide
FT   Protein         28. .225
FT                   /label= Mat_protein
XX
PN   WO9833922-A1.
XX
PD   06-AUG-1998.
XX
PF   02-FEB-1998;   98WO-US002363.
XX
PR   03-FEB-1997;   97US-00792019.
PR   30-JAN-1998;   98US-00016534.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M,  Elliot GS,  Senaldi G,  Sarmiento U;
XX
DR   WPI; 1998-437475/37.
DR   N-PSDB; AAV47510, AAV47511.
XX
PT   Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT   NNT-1 - useful for treatment of neurological and immunological diseases
PT   or inflammation, also as vaccine adjuvant.
XX
PS   Claim 12; Fig 3; 120pp; English.
XX
CC   This is the amino acid sequence of a novel neurotrophic factor,
CC   designated NNT-1, that is a growth factor for neurons and for B or T
CC   cells. It was deduced from isolated cDNA (see AAV47510) and genomic DNA
CC   (see AAV47511) clones. Vectors containing the cDNA or genomic DNA and
CC   host cells are provided for use in the production of NNT-1 polypeptides.
CC   These are used to treat: (i) neurological or immunological diseases,
CC   specifically Alzheimer's, Parkinson's or Huntington's diseases,
CC   amyotrophic lateral sclerosis, Charcot-Marie-Tooth syndrome, peripheral
CC   neuropathy, dystrophy and degeneration of the neural retina, or
CC   conditions characterised by T or B cell defects, e.g. common variable
CC   immunodeficiency (CVID), selective IgA deficiency, hypogammaglobulinaemia
CC   and X-linked agammaglobulinaemia (claimed), but many others disclosed;
CC   and (ii) inflammation. NNT-1 is also able to boost immunoreactivity and
CC   antibody production following vaccination, and, since it inhibits tumour
CC   necrosis factor production, it may also be useful for treating sepsis. In
CC   addition, cells that have been engineered to express NNT-1 can be
CC   implanted, or nucleic acids are delivered in gene therapy vectors
XX
SQ   Sequence 225 AA;

  Query Match             100.0%;  Score 1226;  DB 2;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-119;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 3
AAW94466
ID   AAW94466 standard; protein; 225 AA.
XX
AC   AAW94466;
XX
DT   22-APR-1999  (first entry)
XX
DE   Human cardiotrophin-like cytokine protein.
XX
KW   Human; cardiotrophin-like cytokine; interleukin 6 cytokine family; CLC;
KW   IL-6; diagnosis; detection; immune system-related disorder; cancer;
KW   cardiac disorder; heart failure; hypertension; cancer;
KW   autoimmune disorder; infection.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /label= signal
FT   Protein         28. .225
FT                   /label= Cardiotrophin-like_cytokine
FT   Domain          74. .79
FT                   /label= CD-I
FT                   /note= "conserved domain"
FT   Domain          150. .156
FT                   /label= CD-II
FT                   /note= "conserved domain"
FT   Domain          194. .198
FT                   /label= CD-III
FT                   /note= "conserved domain"
XX
```

PN WO9900415-A1.
XX
PD 07-JAN-1999.
XX
PF 29-JUN-1998; 98WO-US013129.
XX
PR 30-JUN-1997; 97US-0051311P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Shi Y, Ruben SM;
XX
DR WPI; 1999-095678/08.
DR N-PSDB; AAX16161.
XX
PT New isolated cardiotrophin-like cytokine nucleic acid - used to develop
PT products for treating cardiac and immune system disorders, e.g. heart
PT failure, hypertension, cancers, autoimmune disorders and infections.
XX
PS Claim 1; Fig 1; 103pp; English.
XX
CC The present invention relates to a novel cardiotrophin-like cytokine
CC (CLC) protein which is a member of the interleukin 6 (IL-6) cytokine
CC family. The present sequence represents the human CLC protein. The
CC present invention also describes screening methods for identifying
CC agonists and antagonists of CLC activity, as well as methods for
CC detecting cardiac and immune system-related disorders and therapeutic
CC methods for treating cardiac and immune system-related disorders, e.g
CC heart failure, hypertension, cancers, autoimmune disorders and infections
XX
SQ Sequence 225 AA;

Query Match 100.0%; Score 1226; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.8e-119;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
Db 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
Db 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
Db 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
Db 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225

RESULT 4
AAY87813
ID AAY87813 standard; protein; 225 AA.
XX
AC AAY87813;
XX
DT 24-AUG-2000 (first entry)
XX
DE Human NNT-1 protein.
XX
KW NNT-1; human; neurotrophic factor; nootropic; neuroprotective; treatment;
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW retinopathy; immune disorder; hematopoietic disorder.
XX
OS Homo sapiens.
XX
PN US6054294-A.
XX
PD 25-APR-2000.
XX
PF 12-DEC-1997; 97US-00988819.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 2000-338492/29.
DR N-PSDB; AAA39481.
XX
PT New nucleic acids encoding neurotrophic factors useful for stimulating
PT growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS Claim 1c; Fig 3; 42pp; English.
XX
CC This invention describes a novel nucleic acid molecule (I) encoding a
CC novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC which are useful for treating patients in whom various cells of the
CC central, autonomic, or peripheral nervous system have degenerated and/or
CC have been damaged by congenital disease, trauma, mechanical damage,
CC surgery, stroke, ischemia, infection, metabolic disease, nutritional
CC deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC neuropathy induced by diabetes or other metabolic disorders, and/or
CC dystrophies or degeneration of the neural retina such as retinitis
CC pigmentosa, drug-induced retinopathies, stationary forms of night
CC blindness, progressive cone-rod degeneration, immune disorders and
CC hematopoietic disorders. (I) is effective in treating neurological
CC conditions and promotes neuron regeneration. Neural functions are
CC effectively restored in patients suffering from various neurological
CC disorders. This sequence represents the human NNT-1 protein described in
CC the method of the invention
XX
SQ Sequence 225 AA;

Query Match 100.0%; Score 1226; DB 3; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.8e-119;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
Db 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
Db 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
Db 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
Db 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225

RESULT 5
AAG63543
ID AAG63543 standard; protein; 225 AA.
XX
AC AAG63543;
XX
DT 15-OCT-2001 (first entry)
XX
DE Amino acid sequence of a human NNT-1 protein.
XX

KW NNT-1; CLF-1; sCNTFRalpha; nervous system; neuron; nervous system;
KW neuro-muscular function; tumour; immune system; haematopoietic system;
KW reproductive system; liver; skeletal muscle; neurodegenerative disease;
KW amyotrophic lateral sclerosis; Parkinson's disease; Huntingdon's disease;
KW muscular mass; paralysis; cancer; obesity; fertility; endometriosis;
KW blastocyst implantation; thrombosis; retinal disease;
KW retinal pigmentosis.
XX
OS Homo sapiens.
XX
PN WO200155172-A2.
XX
PD 02-AUG-2001.
XX
PF 26-JAN-2001; 2001WO-FR000253.
XX
PR 27-JAN-2000; 2000FR-00001035.
PR 12-OCT-2000; 2000FR-00013089.
XX
PA (FABR) FABRE MEDICAMENT SA PIERRE.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Elson G, Gauchat J, Plun-Favreau H, Chevalier S, Gascan H;
XX
DR WPI; 2001-488773/53.
DR N-PSDB; AAH74484.
XX
PT A complex comprising a NNT-1 protein and a CLF-1 and/or sCNTFRalpha
PT protein useful to treat neurodegenerative disease including Parkinson's
PT and Huntington's, obesity and cancer.
XX
PS Claim 2; Page 58; 67pp; French.
XX
CC The present sequence represents a human NNT-1 protein. The specification
CC describes a complex comprising a NNT-1 protein and a CLF-1 and/or
CC sCNTFRalpha protein. The NNT-1/CLF-1 complex is used to modulate activity
CC of the sCNTFRalpha/gp130/LIFRbeta receptor complex, or to induce
CC phosphorylation of the tyrosine of gp130 and LIFRbeta, particularly where
CC cells expressing the receptor complex are in the central or peripheral
CC nervous system, in neurons implicated in neuro-muscular function or in
CC skeletal muscle. The complex or antibodies are also used to decrease the
CC survival, growth or proliferation of tumour cells or to facilitate the
CC proliferation and/or inhibit differentiation of cells stocks. The complex
CC is also used to modulate activity of the gp130/LIFRbeta receptor or cells
CC expressing that receptor, particularly those cells implicated in the
CC immune, haematopoietic, nervous or reproductive system, the liver or
CC skeletal muscle. Molecules of the invention may be used to prevent or
CC treat neurodegenerative diseases including amyotrophic lateral sclerosis,
CC Parkinson's and Huntington's disease, to repair or regenerate nervous or
CC muscular tissue or to maintain muscular mass in paralysis patients. They
CC may also be used to treat cancer, obesity and associated diseases, and to
CC improve fertility, particularly to avoid endometriosis and/or assist
CC blastocyst implantation, thrombosis, or retinal disease, particular
CC retinal pigmentosis
XX
SQ Sequence 225 AA;

Query Match 100.0%; Score 1226; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.8e-119;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
```

RESULT 6
AAU78176
ID AAU78176 standard; protein; 225 AA.
XX
AC AAU78176;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human novel neurotrophic factor NNT1.
XX
KW Human; NNT1; neurotrophic factor; IgE-related disease;
KW Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW vascular restenosis; rheumatoid arthritis; psoriatic arthritis;
KW inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW inflammatory bowel disease; transplant rejection; reproductive disorder;
KW graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS Homo sapiens.
XX
PN WO200215977-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-US025906.
XX
PR 18-AUG-2000; 2000US-0226436P.
PR 16-AUG-2001; 2001US-00931704.
XX
PA (AMGE-) AMGEN INC.
XX
PI Senaldi G;
XX
DR WPI; 2002-280867/32.
DR N-PSDB; ABK11647.
XX
PT Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT patient, preventing IgE-related disease and treating allergic diseases,
PT involves administering NNT-1 inhibitor to a patient.
XX
PS Claim 2; Fig 3; 63pp; English.
XX
CC The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC modulating IgE levels in a patient, preventing an IgE-related disease,
CC and treating allergic diseases, comprising administering a
CC therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC related disease or susceptibility to an IgE-related disease, by
CC determining the presence or amount of expression of an NNT1 polypeptide
CC encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC occurring variant, and diagnosing an IgE-related disease or
CC susceptibility of an IgE-related disease based on the presence or amount
CC of expression of the polypeptide and a pharmaceutical composition for use
CC in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC inhibitor is useful for preventing and treating IgE-related disease,
CC modulating IgE levels, and treating allergic diseases e.g. Type I
CC allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC asthma, immune diseases and disorders, diseases involving abnormal cell
CC proliferation including cancer, arteriosclerosis and vascular restenosis,
CC diseases and conditions relating to dysfunction of immune system
CC including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC transplant rejection, and graft versus host disease, and reproductive
CC diseases and disorders including infertility, miscarriage, preterm labour
CC and delivery, and endometriosis. The present sequence represents human
CC NNT1

```
XX
SQ   Sequence 225 AA;

  Query Match             100.0%;  Score 1226;  DB 5;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-119;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 7
AAM25831
ID   AAM25831 standard; protein; 253 AA.
XX
AC   AAM25831;
XX
DT   16-OCT-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:1346.
XX
KW   Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW   antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW   antibacterial; endocrine; cardiant; central nervous system; virucide;
KW   anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW   antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW   dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW   neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW   immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW   antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW   cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW   genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW   thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW   allergic rhinitis; diabetes; multiple sclerosis; depression;
KW   Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW   neurological disorder.
XX
OS   Homo sapiens.
XX
PN   WO200153455-A2.
XX
PD   26-JUL-2001.
XX
PF   22-DEC-2000; 2000WO-US035017.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457603/49.
DR   N-PSDB; AAH99772.
XX
PT   Isolated human polynucleotides encoding polypeptides, useful for the
PT   treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS   Claim 20; Page 278; 1217pp; English.
XX
CC   AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC   AAM25963. The proteins can have activities based on the tissues and cells
CC   they are expressed in, such as: antiinflammatory; antirheumatic;
CC   antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC   central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC   cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC   antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic;
CC   antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC   antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC   encoding them can be used in gene therapy, antisense therapy and vaccine
CC   production, The proteins and polynucleotides are useful for screening for
CC   agonists or antagonists of a protein and for the treatment and diagnosis
CC   of disorders associated with the activity of a protein e.g. inflammation,
CC   rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC   neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC   infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC   anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC   osteoporosis, severe combined immunodeficiency, eczema, allergic
CC   rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC   Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC   neurological disorders
XX
SQ   Sequence 253 AA;

  Query Match             100.0%;  Score 1226;  DB 4;  Length 253;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-119;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 88

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 148

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 208

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        209 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 253


RESULT 8
AAE00828
ID   AAE00828 standard; protein; 223 AA.
XX
AC   AAE00828;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human cardiotrophin-like cytokine (CLC) protein.
XX
KW   Human; biologically active complex; haemopoietin receptor; NR6;
KW   cardiotrophin-like cytokine; CLC; therapy; prophylaxis; proliferation;
KW   differentiation; cell survival; neurotrophic activity.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /label= Signal_peptide
FT   Protein         28. .223
FT                   /label= Human_mature_CLC_protein
FT                   /note= "Cardiotrophin-like cytokine"
XX
PN   WO200127157-A1.
XX
PD   19-APR-2001.
XX
```

PF 06-OCT-2000; 2000WO-AU001216.
XX
PR 08-OCT-1999; 99AU-00003327.
PR 12-MAY-2000; 2000AU-00007489.
XX
PA (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
PI Nash A, Jachno KM, Fabri LJ, Reid K, Bartlett PF, Hilton DJ;
PI Nakata Y, Hasegawa M;
XX
DR WPI; 2001-281978/29.
DR N-PSDB; AAD04201.
XX
PT New biologically active complex comprising NR6 and cardiotrophin-like-
PT cytokine, for facilitating proliferation, differentiation and/or survival
PT of a cell.
XX
PS Claim 32; Page 114-115; 123pp; English.
XX
CC The present invention relates to a biologically active complex comprising
CC a haemopoietin receptor, NR6 and cardiotrophin-like cytokine (CLC). The
CC complex is useful in the manufacture of a medicament for the treatment
CC and/or prophylaxis of a subject, as it is involved in facilitating
CC proliferation, differentiation and/or survival of a cell. The complex or
CC its components have neurotrophic activity. The present sequence is human
CC cardiotrophin-like cytokine (CLC) protein
XX
SQ Sequence 223 AA;

Query Match 99.0%; Score 1214; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 8.4e-118;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
Db 1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
Db 61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
Db 121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAH 223
Db 181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAH 223

RESULT 9
ABB11896
ID ABB11896 standard; peptide; 321 AA.
XX
AC ABB11896;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human cardiotrophin-like cytokine homologue, SEQ ID NO:2266.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR N-PSDB; ABA09140.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 20; Page 273; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention
XX
SQ Sequence 321 AA;

Query Match 98.2%; Score 1204; DB 4; Length 321;
Best Local Similarity 99.1%; Pred. No. 1.5e-116;
Matches 221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 62
Db 99 LPTGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 158

```
Qy         63 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 218

Qy        123 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 278

Qy        183 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||
Db        279 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 321
```

RESULT 10
AAM79399
ID AAM79399 standard; protein; 321 AA.
XX
AC AAM79399;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3045.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK52532.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 237; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX

SQ Sequence 321 AA;

Query Match 98.2%; Score 1204; DB 4; Length 321;
Best Local Similarity 99.1%; Pred. No. 1.5e-116;
Matches 221; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 62
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 LPTGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 158

Qy         63 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 218

Qy        123 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 278

Qy        183 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||
Db        279 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 321
```

RESULT 11
AAM78415
ID AAM78415 standard; protein; 260 AA.
XX
AC AAM78415;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1077.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51548.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 3306; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The

```
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 260 AA;

  Query Match             97.6%;  Score 1197;  DB 4;  Length 260;
  Best Local Similarity   98.7%;  Pred. No. 6.1e-116;
  Matches  220;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 62
              | ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         38 LPTGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDPTRYLEHQLRSLAGTYL 97

Qy         63 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 157

Qy        123 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 217

Qy        183 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||
Db        218 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 260


RESULT 12
AAW56142
ID   AAW56142 standard; protein; 225 AA.
XX
AC   AAW56142;
XX
DT   13-JUL-1998  (first entry)
XX
DE   Amino acid sequence of murine neurotrophic factor NNT-1.
XX
KW   Mouse; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW   treatment; neurological disease; degeneration; Parkinson's disease;
KW   amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /note= "signal peptide"
FT   Protein         28. .225
FT                   /note= "mature peptide"
XX
PN   US5741772-A.
XX
PD   21-APR-1998.
XX
PF   03-FEB-1997;   97US-00792019.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 1998-260526/23.
DR   N-PSDB; AAV22654.
XX
PT   Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT   for stimulating growth of motor and sympathetic neurons.
XX
PS   Claim 2; Fig 5; 41pp; English.
XX
CC   The present sequence represents a murine neurotrophic factor, designated
CC   NNT-1, which is capable of stimulating growth of motor or sympathetic
CC   neurons. The NNT-1 protein is useful in the treatment of neurological
CC   diseases characterised by the degeneration and death of particular
CC   classes of neurons. These diseases specifically include Parkinson's
CC   disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke
CC   and various degenerative disorders affecting vision
XX
SQ   Sequence 225 AA;

  Query Match             97.3%;  Score 1193;  DB 2;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 1.3e-115;
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              |||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||||||||| |||||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 13
AAW29716
ID   AAW29716 standard; protein; 225 AA.
XX
AC   AAW29716;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Mouse neurotrophic factor NNT-1.
XX
KW   NNT-1; neurotrophic factor; mouse; antiinflammatory; adjuvant;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   peripheral neurophathy; dystrophy; neural retina degeneration;
KW   common variable immunodeficiency; CVID; selective IgA deficiency;
KW   hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW   therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .27
FT                   /label= Sig_peptide
FT   Protein         28. .225
FT                   /label= Mat_protein
XX
PN   WO9833922-A1.
XX
PD   06-AUG-1998.
XX
PF   02-FEB-1998;   98WO-US002363.
XX
PR   03-FEB-1997;   97US-00792019.
PR   30-JAN-1998;   98US-00016534.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M,  Elliot GS,  Senaldi G,  Sarmiento U;
XX
DR   WPI; 1998-437475/37.
```

DR N-PSDB; AAV47512.
XX
PT Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT NNT-1 - useful for treatment of neurological and immunological diseases
PT or inflammation, also as vaccine adjuvant.
XX
PS Claim 13; Fig 5; 120pp; English.
XX
CC This is the amino acid sequence of a murine neurotrophic factor,
CC designated NNT-1, that is a growth factor for neurons and for B or T
CC cells. It was deduced from isolated NNT-1 cDNA (see AAV47512). Human NNT-
CC 1 (see AAW29715) is also provided. Vectors and host cells for use in the
CC production of human murine recombinant NNT-1 polypeptides. These are used
CC to treat: (i) neurological or immunological diseases, specifically
CC Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic lateral
CC sclerosis, Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy
CC and degeneration of the neural retina, or conditions characterised by T
CC or B cell defects, e.g. common variable immunodeficiency (CVID),
CC selective IgA deficiency, hypogammaglobulinaemia and X-linked
CC agammaglobulinaemia (claimed), but many others disclosed; and (ii)
CC inflammation. NNT-1 is also able to boost immunoreactivity and antibody
CC production following vaccination, and, since it inhibits tumour necrosis
CC factor production, it may also be useful for treating sepsis. In
CC addition, cells that have been engineered to express NNT-1 can be
CC implanted, or nucleic acids are delivered in gene therapy vectors
XX
SQ Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 2; Length 225;
Best Local Similarity 96.9%; Pred. No. 1.3e-115;
Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||| |||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225
```

RESULT 14
AAY87814
ID AAY87814 standard; protein; 225 AA.
XX
AC AAY87814;
XX
DT 24-AUG-2000 (first entry)
XX
DE Murine NNT-1 protein.
XX
KW NNT-1; neurotrophic factor; nootropic; neuroprotective; treatment;
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome; murine;
KW Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW retinopathy; immune disorder; hematopoietic disorder.
XX
OS Mus sp.
XX
PN US6054294-A.
XX
PD 25-APR-2000.
XX
PF 12-DEC-1997; 97US-00988819.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 2000-338492/29.
DR N-PSDB; AAA39483.
XX
PT New nucleic acids encoding neurotrophic factors useful for stimulating
PT growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS Claim 2b; Fig 5; 42pp; English.
XX
CC This invention describes a novel nucleic acid molecule (I) encoding a
CC novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC which are useful for treating patients in whom various cells of the
CC central, autonomic, or peripheral nervous system have degenerated and/or
CC have been damaged by congenital disease, trauma, mechanical damage,
CC surgery, stroke, ischemia, infection, metabolic disease, nutritional
CC deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC neuropathy induced by diabetes or other metabolic disorders, and/or
CC dystrophies or degeneration of the neural retina such as retinitis
CC pigmentosa, drug-induced retinopathies, stationary forms of night
CC blindness, progressive cone-rod degeneration, immune disorders and
CC hematopoietic disorders. (I) is effective in treating neurological
CC conditions and promotes neuron regeneration. Neural functions are
CC effectively restored in patients suffering from various neurological
CC disorders. This sequence represents the murine NNT-1 protein described in
CC the method of the invention
XX
SQ Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 3; Length 225;
Best Local Similarity 96.9%; Pred. No. 1.3e-115;
Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||| |||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225
```

RESULT 15
AAU78177
ID AAU78177 standard; protein; 225 AA.
XX
AC AAU78177;
XX
DT 05-JUN-2002 (first entry)
XX
DE Mouse novel neurotrophic factor NNT1.
XX
KW Mouse; NNT1; neurotrophic factor; IgE-related disease;
KW Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW vascular restenosis; rheumatoid arthritis; psoriatic arthritis;

KW inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW inflammatory bowel disease; transplant rejection; reproductive disorder;
KW graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS Mus sp.
XX
PN WO200215977-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-US025906.
XX
PR 18-AUG-2000; 2000US-0226436P.
PR 16-AUG-2001; 2001US-00931704.
XX
PA (AMGE-) AMGEN INC.
XX
PI Senaldi G;
XX
DR WPI; 2002-280867/32.
DR N-PSDB; ABK11649.
XX
PT Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT patient, preventing IgE-related disease and treating allergic diseases,
PT involves administering NNT-1 inhibitor to a patient.
XX
PS Claim 2; Fig 5; 63pp; English.
XX
CC The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC modulating IgE levels in a patient, preventing an IgE-related disease,
CC and treating allergic diseases, comprising administering a
CC therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC related disease or susceptibility to an IgE-related disease, by
CC determining the presence or amount of expression of an NNT1 polypeptide
CC encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC occurring variant, and diagnosing an IgE-related disease or
CC susceptibility of an IgE-related disease based on the presence or amount
CC of expression of the polypeptide and a pharmaceutical composition for use
CC in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC inhibitor is useful for preventing and treating IgE-related disease,
CC modulating IgE levels, and treating allergic diseases e.g. Type I
CC allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC asthma, immune diseases and disorders, diseases involving abnormal cell
CC proliferation including cancer, arteriosclerosis and vascular restenosis,
CC diseases and conditions relating to dysfunction of immune system
CC including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC transplant rejection, and graft versus host disease, and reproductive
CC diseases and disorders including infertility, miscarriage, preterm labour
CC and delivery, and endometriosis. The present sequence represents Mouse
CC NNT1
XX
SQ Sequence 225 AA;

Query Match 97.3%; Score 1193; DB 5; Length 225;
Best Local Similarity 96.9%; Pred. No. 1.3e-115;
Matches 218; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||| |||||||||||||| | ||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225
```

Search completed: August 10, 2004, 06:45:59
Job time : 51 secs

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:46:33 ; Search time 513 Seconds
(without alignments)
137.580 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1226	100.0	225	9	US-09-931-704-2	Sequence 2, Appli
2	1226	100.0	225	14	US-10-212-793-2	Sequence 2, Appli
3	1226	100.0	225	16	US-10-408-765A-785	Sequence 785, App
4	1226	100.0	253	12	US-10-296-115-1346	Sequence 1346, Ap
5	1204	98.2	321	12	US-10-276-774-2266	Sequence 2266, Ap
6	1193	97.3	225	9	US-09-931-704-5	Sequence 5, Appli
7	885	72.2	164	9	US-09-864-761-40014	Sequence 40014, A
8	124.5	10.2	203	14	US-10-212-793-3	Sequence 3, Appli
9	118.5	9.7	203	9	US-09-896-856-3	Sequence 3, Appli
10	118.5	9.7	203	14	US-10-107-931-3	Sequence 3, Appli
11	118.5	9.7	203	15	US-10-407-303-3	Sequence 3, Appli
12	118.5	9.7	203	16	US-10-722-095-3	Sequence 3, Appli
13	105	8.6	471	12	US-10-282-122A-69553	Sequence 69553, A
14	96.5	7.9	243	9	US-09-810-052-5	Sequence 5, Appli
15	96.5	7.9	243	10	US-09-791-497-8	Sequence 8, Appli
16	96.5	7.9	243	13	US-10-000-776-6	Sequence 6, Appli
17	92.5	7.5	860	14	US-10-156-761-11445	Sequence 11445, A
18	91.5	7.5	201	9	US-09-901-540-3	Sequence 3, Appli
19	91.5	7.5	201	9	US-09-896-856-8	Sequence 8, Appli
20	91.5	7.5	201	9	US-09-901-257-3	Sequence 3, Appli
21	91.5	7.5	201	14	US-10-107-931-8	Sequence 8, Appli
22	91.5	7.5	201	14	US-10-212-793-4	Sequence 4, Appli
23	91.5	7.5	201	15	US-10-407-303-8	Sequence 8, Appli
24	91.5	7.5	201	16	US-10-722-095-8	Sequence 8, Appli
25	91.5	7.5	693	16	US-10-452-858C-93	Sequence 93, Appl
26	91.5	7.5	1182	14	US-10-024-368-6	Sequence 6, Appli
27	91	7.4	195	9	US-09-770-361-5	Sequence 5, Appli
28	91	7.4	195	14	US-10-226-759-5	Sequence 5, Appli
29	91	7.4	195	14	US-10-281-643-5	Sequence 5, Appli
30	91	7.4	195	15	US-10-383-916-5	Sequence 5, Appli
31	91	7.4	242	10	US-09-791-497-2	Sequence 2, Appli
32	91	7.4	242	13	US-10-000-776-2	Sequence 2, Appli
33	90	7.3	232	9	US-09-810-052-2	Sequence 2, Appli
34	89.5	7.3	1313	12	US-10-437-421-22	Sequence 22, Appl
35	88.5	7.2	319	12	US-10-425-114-61252	Sequence 61252, A
36	88.5	7.2	323	12	US-10-425-114-38754	Sequence 38754, A
37	88.5	7.2	1429	12	US-10-029-347-3	Sequence 3, Appli
38	88.5	7.2	1429	14	US-10-028-374-3	Sequence 3, Appli
39	88.5	7.2	1429	14	US-10-183-770-3	Sequence 3, Appli
40	87.5	7.1	1207	14	US-10-024-368-5	Sequence 5, Appli
41	85.5	7.0	499	12	US-10-425-114-45547	Sequence 45547, A
42	85.5	7.0	1022	12	US-10-188-186-26	Sequence 26, Appl
43	85	6.9	727	15	US-10-231-913-98	Sequence 98, Appl
44	84.5	6.9	389	15	US-10-369-493-22941	Sequence 22941, A
45	84.5	6.9	1473	12	US-10-182-822A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-931-704-2
; Sequence 2, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
; APPLICANT: Senaldi, Giorgio
; TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: A-695
; CURRENT APPLICATION NUMBER: US/09/931,704
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/226,436
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-931-704-2

Query Match 100.0%; Score 1226; DB 9; Length 225;
Best Local Similarity 100.0%; Pred. No. 8.8e-115;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 2
US-10-212-793-2
; Sequence 2, Application US/10212793
; Publication No. US20030087395A1
; GENERAL INFORMATION:
;  APPLICANT: Shi et al.
;  TITLE OF INVENTION: Cardiotrophin-Like Cytokine
;  FILE REFERENCE: PF385D1C1
;  CURRENT APPLICATION NUMBER: US/10/212,793
;  CURRENT FILING DATE:  2002-08-07
;  PRIOR APPLICATION NUMBER: US 09/438,299
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: US 09/106,182
;  PRIOR FILING DATE: 1998-06-29
;  PRIOR APPLICATION NUMBER: US 60/051,311
;  PRIOR FILING DATE: 1997-06-30
;  NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 2
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: signal
;   LOCATION: (-27)..(-1)
;   FEATURE:
;   NAME/KEY: peptide
;   LOCATION: (1)..(198)
US-10-212-793-2

  Query Match             100.0%;  Score 1226;  DB 14;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-115;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 3
US-10-408-765A-785
; Sequence 785, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
;  APPLICANT: Ghosh, Soumitra S.
;  APPLICANT:  Fahy, Eoin D.
;  APPLICANT:  Zhang, Bing
;  APPLICANT:  Gibson, Bradford W.
;  APPLICANT:  Taylor, Steven W.
;  APPLICANT:  Glenn, Gary M.
;  APPLICANT:  Warnock, Dale E.
;  TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;  TITLE OF INVENTION:  IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;  FILE REFERENCE: 660088.465
;  CURRENT APPLICATION NUMBER: US/10/408,765A
;  CURRENT FILING DATE:  2003-04-04
;  NUMBER OF SEQ ID NOS: 3077
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 785
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-408-765A-785

  Query Match             100.0%;  Score 1226;  DB 16;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-115;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 4
US-10-296-115-1346
; Sequence 1346, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 784PCT
;  CURRENT APPLICATION NUMBER: US/10/296,115
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: US09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1346
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-296-115-1346

  Query Match             100.0%;  Score 1226;  DB 12;  Length 253;
  Best Local Similarity   100.0%;  Pred. No. 1e-114;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         29 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 88

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 148

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 208

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        209 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 253


RESULT 5
US-10-276-774-2266
; Sequence 2266, Application US/10276774
; Publication No. US20040053245A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  APPLICANT:  Tang, Y, Tom et al
;  TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 21272-030
;  CURRENT APPLICATION NUMBER: US/10/276,774
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: 09/560,875
;  PRIOR FILING DATE: 2000-04-27
;  PRIOR APPLICATION NUMBER: 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  NUMBER OF SEQ ID NOS: 2700
;  SOFTWARE: Custom
; SEQ ID NO 2266
;   LENGTH: 321
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-276-774-2266

  Query Match             98.2%;  Score 1204;  DB 12;  Length 321;
  Best Local Similarity   99.1%;  Pred. No. 2.3e-112;
  Matches  221;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 LRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 62
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 LPTGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYL 158

Qy         63 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 NYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNR 218

Qy        123 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 QAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKM 278

Qy        183 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||
Db        279 DDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 321


RESULT 6
US-09-931-704-5
; Sequence 5, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
;  APPLICANT: Senaldi, Giorgio
;  TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
;  TITLE OF INVENTION:  Inhibitors
;  FILE REFERENCE: A-695
;  CURRENT APPLICATION NUMBER: US/09/931,704
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: US 60/226,436
;  PRIOR FILING DATE: 2000-08-18
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: Murine
US-09-931-704-5

  Query Match             97.3%;  Score 1193;  DB 9;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 1.8e-111;
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              ||||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 7
US-09-864-761-40014
; Sequence 40014, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40014
;   LENGTH: 164
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005849.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
```

```
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
;   OTHER INFORMATION: EST_HUMAN HIT: AI752561.1, EVALUE 3.00e-66
;   OTHER INFORMATION: SWISSPROT HIT: Q63086, EVALUE 8.00e-03
US-09-864-761-40014

  Query Match             72.2%;  Score 885;  DB 9;  Length 164;
  Best Local Similarity   99.4%;  Pred. No. 1.1e-80;
  Matches  163;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         62 LNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLN 121
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db          1 LNYLGPPFNEPDFNPPRLGAKTLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLN 60

Qy        122 RQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQK 120

Qy        182 MDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||
Db        121 MDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 164


RESULT 8
US-10-212-793-3
; Sequence 3, Application US/10212793
; Publication No. US20030087395A1
; GENERAL INFORMATION:
;  APPLICANT: Shi et al.
;  TITLE OF INVENTION: Cardiotrophin-Like Cytokine
;  FILE REFERENCE: PF385D1C1
;  CURRENT APPLICATION NUMBER: US/10/212,793
;  CURRENT FILING DATE:  2002-08-07
;  PRIOR APPLICATION NUMBER: US 09/438,299
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: US 09/106,182
;  PRIOR FILING DATE: 1998-06-29
;  PRIOR APPLICATION NUMBER: US 60/051,311
;  PRIOR FILING DATE: 1997-06-30
;  NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 3
;   LENGTH: 203
;   TYPE: PRT
;   ORGANISM: rattus norvegicus
US-10-212-793-3

  Query Match             10.2%;  Score 124.5;  DB 14;  Length 203;
  Best Local Similarity   28.1%;  Pred. No. 0.00041;
  Matches  50;  Conservative   24;  Mismatches   85;  Indels   19;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |      |   |:   | ||  | |:||||    |       :   | |
Db         27 IRQTHNLARLLTKYADQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSALPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGY----PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     |:|:|:  |  :|   |  | |:    :      |: |:  |  :|
Db        141 AALGAAARGPVPEPV-ATSALFTSNSAAGVFSAKVLGLHVCGLYGEWVSRTEGDLGQL 197


RESULT 9
US-09-896-856-3
; Sequence 3, Application US/09896856
; Patent No. US20020137189A1
;   GENERAL INFORMATION:
;        APPLICANT: Baker, Joffre
;                   Chien, Kenneth
;                   King, Kathleen
;                   Pennica, Diane
;                   Wood, William
;        TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genentech, Inc.
;             STREET: 460 Point San Bruno Blvd
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: WinPatin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/896,856
;             FILING DATE: 29-Jun-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 08/733,850
;             FILING DATE: 18-OCT-1996
;             APPLICATION NUMBER: US 08/471,112
;             FILING DATE: 06-JUN-1995
;             APPLICATION NUMBER: 08/233,609
;             FILING DATE: 25-APR-1994
;             APPLICATION NUMBER: 08/286304
;             FILING DATE: 05-AUG-1994
;             APPLICATION NUMBER: 08/443129
;             FILING DATE: 17-MAY-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Conley, Deirdre L.
;             REGISTRATION NUMBER: 36,487
;             REFERENCE/DOCKET NUMBER: P0894P1D2C1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415/225-2066
;             TELEFAX: 415/952-9881
;             TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 203 amino acids
;             TYPE: Amino Acid
;             TOPOLOGY: Linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-896-856-3

  Query Match              9.7%;  Score 118.5;  DB 9;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 0.0016;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |      |   |:   | ||  | |:||||    |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | |  | :|      |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 10
US-10-107-931-3
; Sequence 3, Application US/10107931
; Publication No. US20030054550A1
;   GENERAL INFORMATION:
;        APPLICANT: Baker, Joffre
;                   Chien, Kenneth
;                   King, Kathleen
;                   Pennica, Diane
;                   Wood, William
;        TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
```

```
;       NUMBER OF SEQUENCES: 8
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genentech, Inc.
;            STREET: 460 Point San Bruno Blvd
;            CITY: South San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94080
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: patin (Genentech)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/107,931
;            FILING DATE: 26-Mar-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/364,187
;            FILING DATE: 30-Jul-1999
;            APPLICATION NUMBER: 08/286,304
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Hasak, Janet E.
;            REGISTRATION NUMBER: 28,616
;            REFERENCE/DOCKET NUMBER: 894P1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415/225-1896
;            TELEFAX: 415/952-9881
;            TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 203 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-107-931-3

  Query Match             9.7%;  Score 118.5;  DB 14;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 0.0016;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                || |:  | | |   |  : |: |       | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | |  | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 11
US-10-407-303-3
; Sequence 3, Application US/10407303
; Publication No. US20040006018A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Joffre
;  APPLICANT:  Chien, Kenneth
;  APPLICANT:  King, Kathleen
;  APPLICANT:  Pennica, Diane
;  APPLICANT:  Wood, William
;  TITLE OF INVENTION: Methods of Stimulating the Proliferation and
;  TITLE OF INVENTION:  Differentiation of Satellite Cells Into Myoblasts by
;  TITLE OF INVENTION:  Contact with CT-1
;  FILE REFERENCE: P0994R1D2C1
;  CURRENT APPLICATION NUMBER: US/10/407,303
;  CURRENT FILING DATE:  2003-04-03
;  PRIOR APPLICATION NUMBER: US 09/724,772
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: US 08/797,014
;  PRIOR FILING DATE: 1997-02-07
;  PRIOR APPLICATION NUMBER: US 60/049,998
;  PRIOR FILING DATE: 1996-02-14
;  NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 3
;   LENGTH: 203
;   TYPE: PRT
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: Amino Acid
;   LOCATION: Full -
;   OTHER INFORMATION:
US-10-407-303-3

  Query Match             9.7%;  Score 118.5;  DB 15;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 0.0016;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                || |:  | | |   |  : |: |       | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | |  | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 12
US-10-722-095-3
; Sequence 3, Application US/10722095
; Publication No. US20040116678A1
; GENERAL INFORMATION:
;  APPLICANT: Baker, Jeffre
;  APPLICANT:  Chien, Kenneth
;  APPLICANT:  King, Kathleen
;  APPLICANT:  Pennica, Diane
;  APPLICANT:  Wood, William
;  TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
;  FILE REFERENCE: P0894P1D2C6
;  CURRENT APPLICATION NUMBER: US/10/722,095
;  CURRENT FILING DATE:  2003-11-24
;  PRIOR APPLICATION NUMBER: US 09/896,856
;  PRIOR FILING DATE: 2001-06-29
;  PRIOR APPLICATION NUMBER: US 09/033,114
;  PRIOR FILING DATE: 1998-03-02
;  PRIOR APPLICATION NUMBER: US 08/733,850
;  PRIOR FILING DATE: 1996-10-18
;  PRIOR APPLICATION NUMBER: US 08/443,129
;  PRIOR FILING DATE: 1995-05-17
;  PRIOR APPLICATION NUMBER: US 08/286,304
;  PRIOR FILING DATE: 1994-08-05
;  PRIOR APPLICATION NUMBER: US 08/233,609
;  PRIOR FILING DATE: 1994-04-25
;  NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 3
;   LENGTH: 203
;   TYPE: PRT
;   ORGANISM: Mus Musculus
US-10-722-095-3

  Query Match             9.7%;  Score 118.5;  DB 16;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 0.0016;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83
```

```
Qy          97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
Db          84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy         152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
Db         141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197
```

RESULT 13
US-10-282-122A-69553
; Sequence 69553, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69553
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69553

Query Match 8.6%; Score 105; DB 12; Length 471;
Best Local Similarity 25.0%; Pred. No. 0.11;
Matches 53; Conservative 15; Mismatches 98; Indels 46; Gaps 6;

```
Qy           1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
Db         182 LELREGDPAALLELAAELAATLPLMPC----------DISKAERGYRLLDASSYSLS-- 228

Qy          61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSH-------- 112
Db         229 ------------LPAPTLNAETPLDDAYSALAWHLLGSSQRLAEQYRFNGHWRLLQDWV 275

Qy         113 -LLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTP 171
Db         276 EMLGELRALTSSLGQAAPRTSTAQLRTALDALLED--------WRPLVQ--AGQEDADVR 325

Qy         172 GPAHSDFLQKMDDF-WLLKELQTWLWRSAKDF 202
Db         326 GAAHEQFLEELQDTRWGEFSLNTSRWLLARSW 357
```

RESULT 14
US-09-810-052-5
; Sequence 5, Application US/09810052
; Patent No. US20020009775A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: HELICAL PROTEIN ZALPHA51
; FILE REFERENCE: 00-24
; CURRENT APPLICATION NUMBER: US/09/810,052
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,410
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/199,443
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-810-052-5

Query Match 7.9%; Score 96.5; DB 9; Length 243;
Best Local Similarity 23.7%; Pred. No. 0.34;
Matches 58; Conservative 26; Mismatches 104; Indels 57; Gaps 9;

```
Qy           1 MDLRAGD-SWGMLACLCTVL------WHLPAVPALNRTGDPGPGPSIQK-------TYDL 46
Db           1 MGQTAGDLGWRLSLLLLPLLLVQAGVWGFPRPPG-------RPQLSLQELRREFTVSLHL 53

Qy          47 TRYLEHQLRSLAGTYLN--------YLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLN 98
Db          54 ARKLLSEVRGQAHRFAESHLPGVNLYLLP-----------LG-EQLPDVSLTFQAWRRLS 101

Qy          99 DKLRLTQNYEAYSHLLCYLRGLNRQAATAELRR-SLAHFCTSLQGLLGSIAGVMAALGYP 157
Db         102 DPERLCFISTTLQPFHAPLGGLGTQGRWTNMERMQLWAMRLDLRDLQRHLRFQVLAAGFN 161

Qy         158 LPQP---------------LPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDF 202
Db         162 LPEEEEEEEEEEEEERKGLLPGALGSALQGPAQVSWPQLLSTYRLLHSLELVLSRAVREL 221

Qy         203 NRLKK 207
Db         222 LLLSK 226
```

RESULT 15
US-09-791-497-8
; Sequence 8, Application US/09791497
; Publication No. US20030008343A1
; GENERAL INFORMATION:
; APPLICANT: Timans, Jacqueline C.
; APPLICANT: Kastelein, Robert A.
; APPLICANT: Bazan, J. Fernando
; APPLICANT: Pflanz, Stefan
; TITLE OF INVENTION: Mammalian Cytokines; Related Reagents
; FILE REFERENCE: DX01040K2
; CURRENT APPLICATION NUMBER: US/09/791,497
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 09/627,897
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/146,581
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/147,763

```
;  PRIOR FILING DATE: 1999-08-06
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 243
;   TYPE: PRT
;   ORGANISM: primate; surmised Homo sapiens
US-09-791-497-8

  Query Match             7.9%;  Score 96.5;  DB 10;  Length 243;
  Best Local Similarity   23.7%;  Pred. No. 0.34;
  Matches   58;  Conservative   26;  Mismatches  104;  Indels   57;  Gaps    9;

Qy          1 MDLRAGD-SWGMLACLCTVL------WHLPAVPALNRTGDPGPGPSIQK-------TYDL 46
              |   |||  | :   |  :|      | |  |          |  |:|:       :  |
Db          1 MGQTAGDLGWRLSLLLLPLLLVQAGVWGFPRPPG-------RPQLSLQELRREFTVSLHL 53

Qy         47 TRYLEHQLRSLAGTYLN--------YLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLN 98
               | |  ::|  |  :          || |           || | ||  ::  : || |:
Db         54 ARKLLSEVRGQAHRFAESHLPGVNLYLLP----------LG-EQLPDVSLTFQAWRRLS 101

Qy         99 DKLRLTQNYEAYSHLLCYLRGLNRQAATAELRR-SLAHFCTSLQGLLGSIAGVMAALGYP 157
              |  ||             | ||  |     : |  |      |: |   :   : | |:
Db        102 DPERLCFISTTLQPFHAPLGGLGTQGRWTNMERMQLWAMRLDLRDLQRHLRFQVLAAGFN 161

Qy        158 LPQP--------------LPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDF 202
              ||:               |||   :   |||   : | :  : ||  |:  | |: ::
Db        162 LPEEEEEEEEEEEEERKGLLPGALGSALQGPAQVSWPQLLSTYRLLHSLELVLSRAVREL 221

Qy        203 NRLKK 207
              | |
Db        222 LLLSK 226


Search completed: August 10, 2004, 07:12:34
Job time : 515 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 06:43:28 ; Search time 15 Seconds
(without alignments)
774.390 Million cell updates/sec

Title: US-09-931-704-2
Perfect score: 1226
Sequence: 1 MDLRAGDSWGMLACLCTVLW.....KKKMQPPAAAVTLHLGAHGF 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1226	100.0	225	1	US-08-792-019B-2	Sequence 2, Appli
2	1226	100.0	225	3	US-09-106-182-2	Sequence 2, Appli
3	1226	100.0	225	3	US-08-988-819-2	Sequence 2, Appli
4	1226	100.0	225	3	US-09-016-534-2	Sequence 2, Appli
5	1193	97.3	225	1	US-08-792-019B-5	Sequence 5, Appli
6	1193	97.3	225	3	US-08-988-819-5	Sequence 5, Appli
7	1193	97.3	225	3	US-09-016-534-5	Sequence 5, Appli
8	124.5	10.2	203	3	US-09-106-182-3	Sequence 3, Appli
9	118.5	9.7	203	1	US-08-233-609-3	Sequence 3, Appli
10	118.5	9.7	203	1	US-08-444-083-3	Sequence 3, Appli
11	118.5	9.7	203	1	US-08-286-304-3	Sequence 3, Appli
12	118.5	9.7	203	1	US-08-442-745-3	Sequence 3, Appli
13	118.5	9.7	203	1	US-08-443-129-3	Sequence 3, Appli
14	118.5	9.7	203	1	US-08-443-952-3	Sequence 3, Appli
15	118.5	9.7	203	1	US-08-443-130-3	Sequence 3, Appli
16	118.5	9.7	203	3	US-08-898-911-3	Sequence 3, Appli
17	118.5	9.7	203	5	PCT-US95-04467-3	Sequence 3, Appli
18	91.5	7.5	201	1	US-08-444-083-8	Sequence 8, Appli
19	91.5	7.5	201	1	US-08-286-304-8	Sequence 8, Appli
20	91.5	7.5	201	1	US-08-442-745-8	Sequence 8, Appli
21	91.5	7.5	201	1	US-08-443-129-8	Sequence 8, Appli
22	91.5	7.5	201	1	US-08-443-952-8	Sequence 8, Appli
23	91.5	7.5	201	1	US-08-443-130-8	Sequence 8, Appli
24	91.5	7.5	201	1	US-08-792-019B-11	Sequence 11, Appl
25	91.5	7.5	201	3	US-09-106-182-4	Sequence 4, Appli
26	91.5	7.5	201	3	US-08-988-819-11	Sequence 11, Appl
27	91.5	7.5	201	3	US-08-898-911-8	Sequence 8, Appli
28	91.5	7.5	201	3	US-09-016-534-11	Sequence 11, Appl
29	91.5	7.5	201	4	US-09-648-183-3	Sequence 3, Appli
30	91.5	7.5	201	5	PCT-US95-04467-8	Sequence 8, Appli
31	91.5	7.5	1182	4	US-09-287-354-6	Sequence 6, Appli
32	91	7.4	195	1	US-07-959-284-5	Sequence 5, Appli
33	91	7.4	195	2	US-08-308-736A-5	Sequence 5, Appli
34	91	7.4	195	4	US-08-645-107A-5	Sequence 5, Appli
35	91	7.4	195	4	US-09-197-349-5	Sequence 5, Appli
36	91	7.4	195	4	US-09-031-693-5	Sequence 5, Appli
37	91	7.4	195	4	US-09-454-380-5	Sequence 5, Appli
38	91	7.4	195	5	PCT-US92-08258-2	Sequence 2, Appli
39	91	7.4	195	5	PCT-US93-09649A-5	Sequence 5, Appli
40	91	7.4	195	5	PCT-US93-09649-5	Sequence 5, Appli
41	89	7.3	737	4	US-09-252-991A-22539	Sequence 22539, A
42	87.5	7.1	1207	4	US-09-287-354-5	Sequence 5, Appli
43	85.5	7.0	560	2	US-08-756-317-11	Sequence 11, Appl
44	84.5	6.9	426	4	US-09-252-991A-29288	Sequence 29288, A
45	84	6.9	200	3	US-08-949-155-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-792-019B-2
; Sequence 2, Application US/08792019B
; Patent No. 5741772
; GENERAL INFORMATION:
; APPLICANT: CHANG, MING-SHI
; TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DEHAVILLAND DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,019B
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-442
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-792-019B-2

Query Match 100.0%; Score 1226; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.8e-129;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 2
US-09-106-182-2
; Sequence 2, Application US/09106182
; Patent No. 6046035
;  GENERAL INFORMATION:
;    APPLICANT:  Shi, Yanggu
;    APPLICANT:  Ruben, Steve
;    TITLE OF INVENTION:  Cardiotrophin-Like Cytokine
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/106,182
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/051,053
;      FILING DATE:  30-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF385
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-106-182-2

  Query Match             100.0%;  Score 1226;  DB 3;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-129;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 3
US-08-988-819-2
; Sequence 2, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-988-819-2

  Query Match             100.0%;  Score 1226;  DB 3;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-129;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 4
US-09-016-534-2
; Sequence 2, Application US/09016534
; Patent No. 6143874
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    APPLICANT:  ELLIOTT, GARY S.
;    APPLICANT:  SARMIENTO, ULLA
;    APPLICANT:  SENALDI, GIORGIO
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER
;      CITY:  THOUSAND OAKS
```

```
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,534
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442B
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-016-534-2

  Query Match             100.0%;  Score 1226;  DB 3;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-129;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225


RESULT 5
US-08-792-019B-5
; Sequence 5, Application US/08792019B
; Patent No. 5741772
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  1840 DEHAVILLAND DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/792,019B
;      FILING DATE:  03-FEB-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-792-019B-5

  Query Match             97.3%;  Score 1193;  DB 1;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 1.4e-125;
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              |||||||||||||||||||||||||||||||||||| |||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              ||||||||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 6
US-08-988-819-5
; Sequence 5, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-988-819-5

  Query Match             97.3%;  Score 1193;  DB 3;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 1.4e-125;
```

```
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||| |||||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 7
US-09-016-534-5
; Sequence 5, Application US/09016534
; Patent No. 6143874
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    APPLICANT:  ELLIOTT, GARY S.
;    APPLICANT:  SARMIENTO, ULLA
;    APPLICANT:  SENALDI, GIORGIO
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,534
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442B
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-016-534-5

  Query Match             97.3%;  Score 1193;  DB 3;  Length 225;
  Best Local Similarity   96.9%;  Pred. No. 1.4e-125;
  Matches  218;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGT 60

Qy         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGL 120
              ||||||||||||||||||||||||||||:|||||||||:||||||||||||||||||||||
Db         61 YLNYLGPPFNEPDFNPPRLGAETLPRATVNLEVWRSLNDRLRLTQNYEAYSHLLCYLRGL 120

Qy        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQ 180
              ||||||||||||||||||||||||||||||||||||| |||||||||||||| | ||||||||||
Db        121 NRQAATAELRRSLAHFCTSLQGLLGSIAGVMATLGYPLPQPLPGTEPAWAPGPAHSDFLQ 180

Qy        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF 225
              |||||||||||||||||||||||||||||:||||| ||||
Db        181 KMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAASVTLHLEAHGF 225


RESULT 8
US-09-106-182-3
; Sequence 3, Application US/09106182
; Patent No. 6046035
;  GENERAL INFORMATION:
;    APPLICANT:  Shi, Yanggu
;    APPLICANT:  Ruben, Steve
;    TITLE OF INVENTION:  Cardiotrophin-Like Cytokine
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/106,182
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/051,053
;      FILING DATE:  30-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF385
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-106-182-3

  Query Match             10.2%;  Score 124.5;  DB 3;  Length 203;
  Best Local Similarity   28.1%;  Pred. No. 8e-06;
  Matches   50;  Conservative   24;  Mismatches   85;  Indels   19;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |      :   | |
Db         27 IRQTHNLARLLTKYADQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |    |  : |: |        | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSALPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGY----PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||      |:|:|:  |   :|    |   |  |:    :        |: |:  |  :|
Db        141 AALGAAARGPVPEPV-ATSALFTSNSAAGVFSAKVLGLHVCGLYGEWVSRTEGDLGQL 197


RESULT 9
US-08-233-609-3
; Sequence 3, Application US/08233609
```

```
; Patent No. 5534615
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,609
;      FILING DATE:  25-APR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hasak, Janet E.
;      REGISTRATION NUMBER:  28,616
;      REFERENCE/DOCKET NUMBER:  894
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1896
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-233-609-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
                |::|::| | |       |   |:   | ||  | |:||||   |       :   | ||
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                  || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
                ||||     | | | |  | :|       |  |:  | :      |: |:  | :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 10
US-08-444-083-3
; Sequence 3; Application US/08444083
; Patent No. 5571675
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,083
;      FILING DATE:  17-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286304
;      FILING DATE:  05-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-444-083-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
                |::|::| | |       |   |:   | ||  | |:||||   |       :   | ||
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                  || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
                ||||     | | | |  | :|       |  |:  | :      |: |:  | :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 11
US-08-286-304-3
; Sequence 3, Application US/08286304
; Patent No. 5571893
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,304
;      FILING DATE:  05-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hasak, Janet E.
;      REGISTRATION NUMBER:  28,616
;      REFERENCE/DOCKET NUMBER:  894P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-1896
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-286-304-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                || |:  | | |   |  : |: |        | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 12
US-08-442-745-3
; Sequence 3, Application US/08442745
; Patent No. 5624806
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/442,745
;      FILING DATE:  17-may-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286304
;      FILING DATE:  05-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-442-745-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                || |:  | | |   |  : |: |        | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 13
US-08-443-129-3
; Sequence 3, Application US/08443129
; Patent No. 5627073
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/443,129
;      FILING DATE:  17-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  304
;      FILING DATE:  (null)
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-443-129-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||      | | | |  | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 14
US-08-443-952-3
; Sequence 3, Application US/08443952
; Patent No. 5679545
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/443,952
;      FILING DATE:  17-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286304
;      FILING DATE:  05-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-443-952-3

  Query Match             9.7%;  Score 118.5;  DB 1;  Length 203;
  Best Local Similarity   28.8%;  Pred. No. 3.8e-05;
  Matches   51;  Conservative   21;  Mismatches   88;  Indels   17;  Gaps    5;

Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |       :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |         | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||      | | | |  | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197


RESULT 15
US-08-443-130-3
; Sequence 3, Application US/08443130
; Patent No. 5723585
;  GENERAL INFORMATION:
;    APPLICANT:  Baker, Joffre
;    APPLICANT:  Chien, Kenneth
;    APPLICANT:  King, Kathleen
;    APPLICANT:  Pennica, Diane
;    APPLICANT:  Wood, William
;    TITLE OF INVENTION:  Cardiac Hypertrophy Factor and Uses
;    TITLE OF INVENTION:  Therefor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/443,130
;      FILING DATE:  17-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/233609
;      FILING DATE:  25-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/286304
;      FILING DATE:  05-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  894P1D3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  203 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

US-08-443-130-3

Query Match 9.7%; Score 118.5; DB 1; Length 203;
Best Local Similarity 28.8%; Pred. No. 3.8e-05;
Matches 51; Conservative 21; Mismatches 88; Indels 17; Gaps 5;

```
Qy         40 IQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRL---GAETLPRATVDLEVWRS 96
              |::|::| | |       |   |:   | ||  | |:||||   |      :   | |
Db         27 IRQTHNLARLLTKYAEQLLEEYVQQQGEPFGLPGFSPPRLPLAGLSGPAPSHAGLPV--- 83

Qy         97 LNDKLRLTQNYEAYSHLLCYLRGLNRQAA-----TAELRRSLAHFCTSLQGLLGSIAGVM 151
                 || |:  | | |   |  : |: |        | |||      :: |  ::  |:
Db         84 ---SERLRQDAAALSVLPALLDAVRRRQAELNPRAPRLLRSLEDAARQVRALGAAVETVL 140

Qy        152 AALGYPL--PQPLPGTEPT-WTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRL 205
              ||||     | | | |  | :|       |  |:  | :      |: |:  |  :|
Db        141 AALGAAARGPGPEPVTVATLFTANSTAGIFSAKVLGFHVCGLYGEWVSRTEGDLGQL 197
```

Search completed: August 10, 2004, 06:48:27
Job time : 17 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 12:20:54 ; Search time 19013.8 Seconds
(without alignments)
11596.087 Million cell updates/sec

Title: US-09-931-704-3
Perfect score: 5087
Sequence: 1 aacctgcgagtgggcctggc.....cctttgctaagtcttcctca 5087

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		5087	100.0	5087	6	AR002596	AR002596 Sequence
2		5087	100.0	5087	6	AX392088	AX392088 Sequence
3		5087	100.0	5087	6	BD195448	BD195448 The neuro
4		5087	100.0	5087	9	AF176912	AF176912 Homo sapi
c 5		4910.4	96.5	168565	9	AP003419	AP003419 Homo sapi
c 6		4908.8	96.5	169144	2	AC005849	AC005849 Homo sapi
7		4605	90.5	135116	2	AP002437	AP002437 Homo sapi
8		1709.4	33.6	246093	2	AC140073	AC140073 Mus muscu
9		1709.4	33.6	269155	10	AC109138	AC109138 Mus muscu
c 10		1632.6	32.1	258710	2	AC135823	AC135823 Rattus no
11		1463.4	28.8	1710	6	BD132824	BD132824 Cardiotro
12		1463.4	28.8	1736	9	BC012939	BC012939 Homo sapi
13		1461.4	28.7	1689	9	AF172854	AF172854 Homo sapi
14		525.4	10.3	881	6	AX205024	AX205024 Sequence
15		525.4	10.3	881	6	AX205042	AX205042 Sequence
16		523.4	10.3	797	6	AR002595	AR002595 Sequence
17		523.4	10.3	797	6	AX392086	AX392086 Sequence
18		523.4	10.3	797	6	BD195447	BD195447 The neuro
19		523.4	10.3	797	9	AF176911	AF176911 Homo sapi
20		495.4	9.7	680	9	AY049779	AY049779 Homo sapi
21		492.8	9.7	1692	6	AX205060	AX205060 Sequence
22		429.6	8.4	819	6	AR002597	AR002597 Sequence
23		429.6	8.4	819	6	AX392089	AX392089 Sequence
24		429.6	8.4	819	6	BD195449	BD195449 The neuro
25		429.6	8.4	819	10	AF176913	AF176913 Mus muscu
26		330	6.5	342	6	BD073715	BD073715 5'EST of
27		329	6.5	340	6	AX903004	AX903004 Sequence
28		329	6.5	340	6	BD038537	BD038537 Sequence
29		324.6	6.4	396	6	BD132825	BD132825 Cardiotro
30		165.2	3.2	283	6	AX202145	AX202145 Sequence
c 31		148.8	2.9	63347	2	AC110526	AC110526 Mus muscu
32		99	1.9	7218	6	I66494	I66494 Sequence 14
33		82.8	1.6	179145	9	AC009785	AC009785 Homo sapi
34		78	1.5	660	11	BV032921	BV032921 S212P6019
35		78	1.5	248948	2	AC142413	AC142413 Mus muscu
c 36		78	1.5	254982	2	AC102125	AC102125 Mus muscu
c 37		74.8	1.5	233361	2	BX510356	BX510356 Danio rer
38		74.2	1.5	245967	2	AC112091	AC112091 Rattus no
39		73.8	1.5	203982	9	AC067854	AC067854 Homo sapi
40		73.2	1.4	169873	2	BX324116	BX324116 Danio rer
c 41		73.2	1.4	195690	10	AC098726	AC098726 Mus muscu
c 42		72.2	1.4	99419	2	AC140090	AC140090 Ornithorh
43		72.2	1.4	120874	2	AC139116	AC139116 Ornithorh
44		70.6	1.4	95097	9	AL133232	AL133232 Human DNA
45		70.6	1.4	163515	2	AC144753	AC144753 Ornithorh

ALIGNMENTS

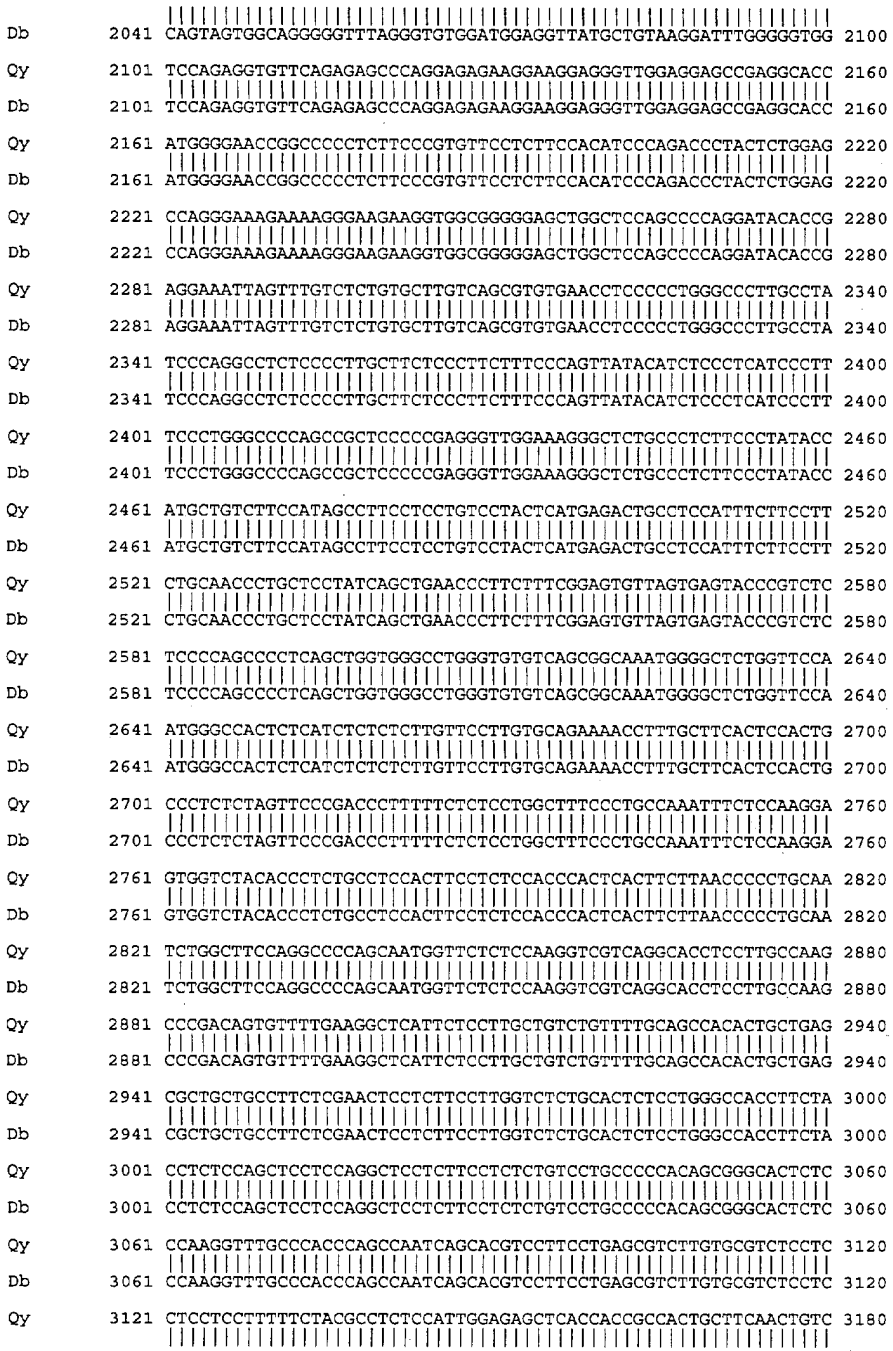
RESULT 1
AR002596
LOCUS AR002596 5087 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5741772.
ACCESSION AR002596
VERSION AR002596.1 GI:3964150
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 5087)
AUTHORS Chang,M.-s.
TITLE Neurotrophic factor NNT-1
JOURNAL Patent: US 5741772-A 3 21-APR-1998;
FEATURES Location/Qualifiers

source 1. .5087
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 5087; DB 6; Length 5087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780

Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900

Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960

Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020

Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080

Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140

Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200

Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260

Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320

Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380

Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440

Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Db       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500

Qy       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Db       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560

Qy       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Db       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620

Qy       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Db       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680

Qy       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Db       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740

Qy       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800

Qy       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860

Qy       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920

Qy       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980

Qy       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040

Qy       2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
```


Db 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Db 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy 4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Db 4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy 4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Db 4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy 4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Db 4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260

```
Qy       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320

Qy       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380

Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440

Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500

Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 2
AX392088

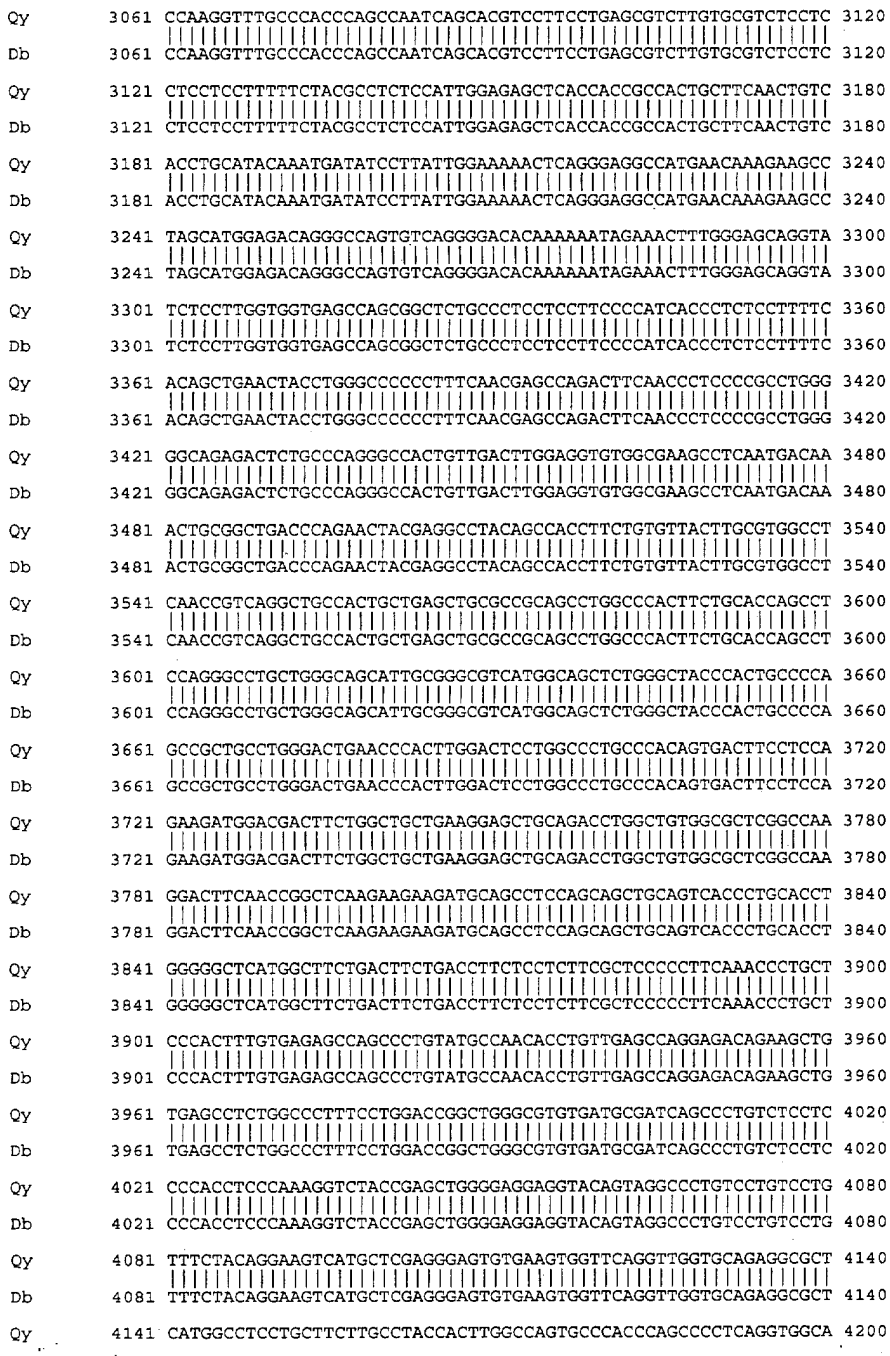
```
LOCUS       AX392088                5087 bp    DNA     linear   PAT 23-MAR-2002
DEFINITION  Sequence 3 from Patent WO0215977.
ACCESSION   AX392088
VERSION     AX392088.1  GI:19700576
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Senaldi,G.
  TITLE     Methods and compositions for treating ige-related disease using nn
            t-1 inhibitors
  JOURNAL   Patent: WO 0215977-A 3 28-FEB-2002;
            Amgen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .5087
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    137. .138
                     /note="INTERVENING UNSEQUENCED REGION OF >1KB"
ORIGIN
```

Query Match 100.0%; Score 5087; DB 6; Length 5087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780

Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

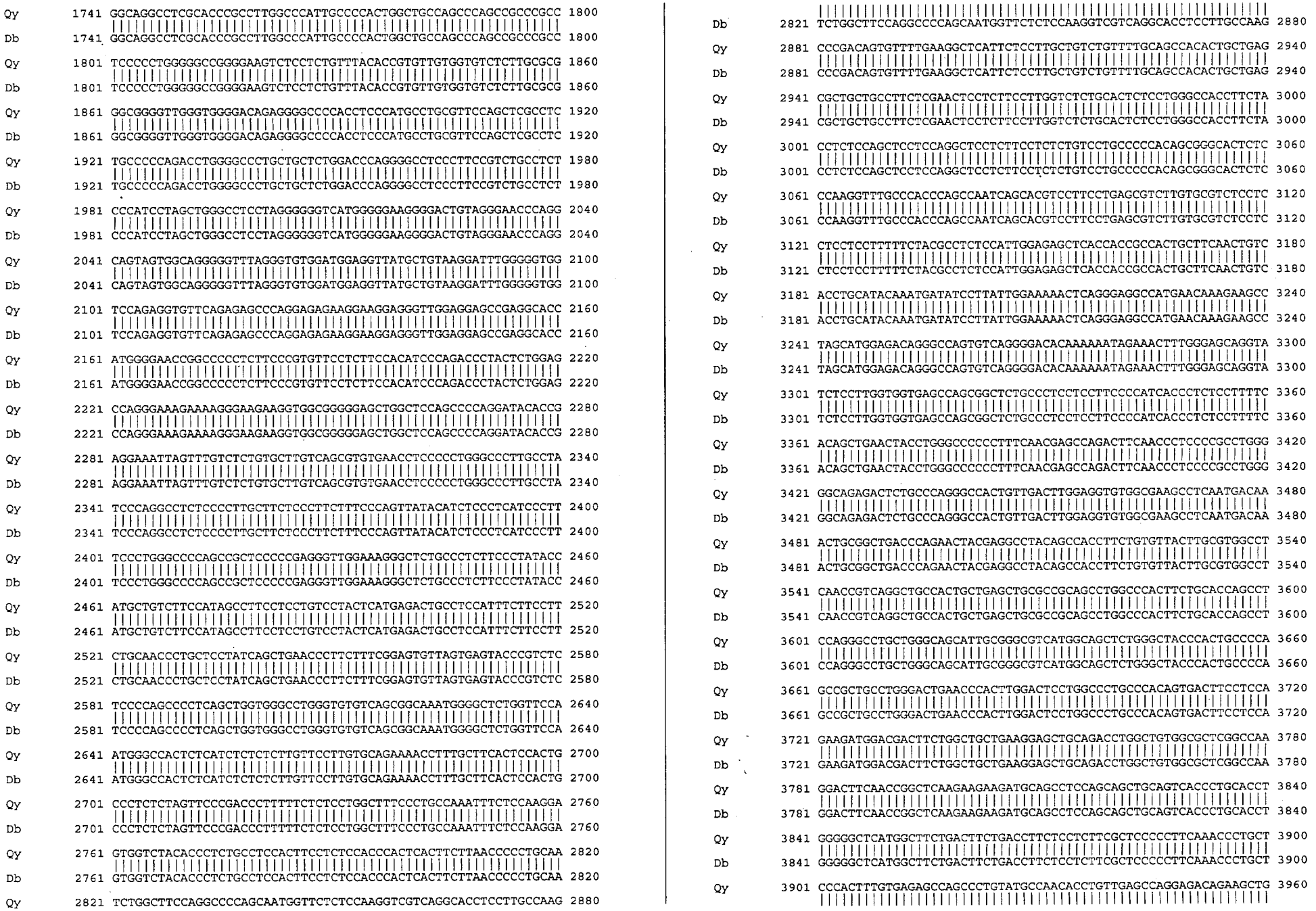
```
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Db       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Qy       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Db       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Qy       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Db       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Qy       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Db       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Qy       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Db       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Qy       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Qy       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Qy       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Qy       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
```

```
Qy       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy       2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db       2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy       2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db       2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy       2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db       2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy       2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db       2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy       2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db       2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy       2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db       2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy       2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db       2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy       2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db       2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy       2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db       2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy       2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db       2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy       2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db       2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy       2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db       2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy       2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db       2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy       2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db       2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy       2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db       2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy       2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db       2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy       3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db       3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
```



RESULT 3
BD195448
LOCUS BD195448 5087 bp DNA linear PAT 17-JUL-2003
DEFINITION The neurotrophic factor NNT-1.
ACCESSION BD195448
VERSION BD195448.1 GI:33005218
KEYWORDS JP 2002514067-A/2.
SOURCE unidentified
ORGANISM unidentified

```
            unclassified.
REFERENCE   1  (bases 1 to 5087)
  AUTHORS   Chang,M.S., Elliot,G.S., Senaldi,G. and Sarmiento,U.
  TITLE     The neurotrophic factor NNT-1
  JOURNAL   Patent: JP 2002514067-A 2 14-MAY-2002;
            AMGEN INC
COMMENT     OS   Unidentified
            PN   JP 2002514067-A/2
            PD   14-MAY-2002
            PF   02-FEB-1998 JP 1998533258
            PR   03-FEB-1997 US    08/792019,30-JAN-1998 US    09/016534 PI
            MING SHI CHANG,GARY S ELLIOT,GIORGI SENALDI,ULLA SARMIENTO PC
            C12N15/18,C07K14/475,C07K16/22,C12N1/21,A61K38/18,A01K67/027 CC
            Strandedness: Single;
            CC   Topology: Linear;
            CC   /product= 'INTERVENING UNSEQUENCED REGION OF >1KB' FH   Key
                 Location/Qualifiers
            FT   source          1. .5087
            FT                   /organism='Unidentified'.
FEATURES                Location/Qualifiers
     source             1. .5087
                        /organism="unidentified"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:32644"
ORIGIN

  Query Match             100.0%;  Score 5087;  DB 6;  Length 5087;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780

Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900

Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960

Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020

Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080

Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140

Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200

Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260

Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320

Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380

Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440

Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500

Qy       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560

Qy       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620

Qy       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680

Qy       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
```


Qy 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Qy 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Qy 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Qy 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Qy 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960

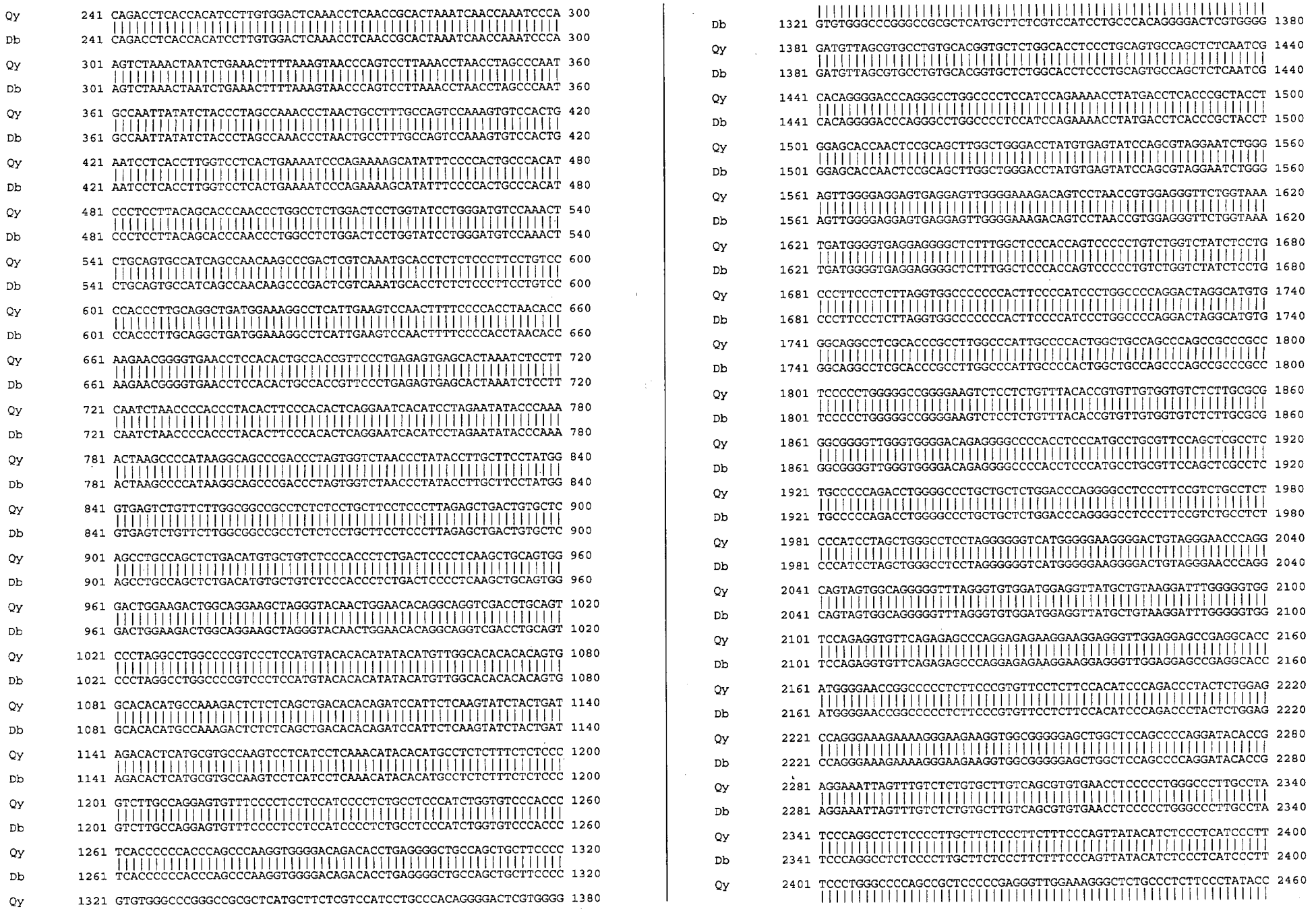
```
Db       3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Db       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Db       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Db       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Qy       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Db       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Qy       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Db       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
```

```
Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 4
AF176912

```
LOCUS       AF176912                5087 bp    DNA     linear   PRI 04-OCT-1999
DEFINITION  Homo sapiens neurotrophin-1/B-cell stimulating factor-3 gene,
            complete cds.
ACCESSION   AF176912
VERSION     AF176912.1  GI:6007642
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5087)
  AUTHORS   Senaldi,G., Varnum,B.C., Sarmiento,U., Starnes,C., Lile,J.,
            Scully,S., Guo,J., Elliott,G., McNinch,J., Shaklee,C.L.,
            Freeman,D., Manu,F., Simonet,W.S., Boone,T. and Chang,M.S.
  TITLE     Novel neurotrophin-1/B cell-stimulating factor-3: a cytokine of the
            IL-6 family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11458-11463 (1999)
  MEDLINE   99432254
   PUBMED   10500198
REFERENCE   2  (bases 1 to 5087)
  AUTHORS   Senaldi,G., Varnum,B., Sarmiento,U., Lile,J., Starnes,C.,
            Scully,S., Guo,J., Elliott,G., McNinch,J., Freeman,D., Shaklee,C.,
            Manu,F., Simonet,S., Boone,T. and Chang,M.-S.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-AUG-1999) Amgen, Inc., One Amgen Center Drive,
            Thousand Oaks, CA 91320, USA
FEATURES             Location/Qualifiers
     source          1. .5087
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     mRNA            join(<120. .135,1368. .1534,3365. .>3859)
                     /product="neurotrophin-1/B-cell stimulating factor-3"
     CDS             join(120. .135,1368. .1534,3365. .3859)
                     /codon_start=1
                     /product="neurotrophin-1/B-cell stimulating factor-3"
                     /protein_id="AAF00992.1"
                     /db_xref="GI:6007643"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN
```

```
  Query Match           100.0%;  Score 5087;  DB 9;  Length 5087;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
```



```
Db      2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540

Qy      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy      3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db      3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy      3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db      3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy      3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db      3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy      3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db      3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy      3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Db      3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy      3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db      3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy      4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db      4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy      4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Db      4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy      4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Db      4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy      4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Db      4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Qy      4261 TTGGCTGCCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Db      4261 TTGGCTGCCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Qy      4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Db      4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Qy      4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Db      4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Qy      4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Db      4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Qy      4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
Db      4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
Qy      4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
Db      4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
```

```
Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

```
RESULT 5
AP003419/c
LOCUS       AP003419               168565 bp    DNA     linear   PRI 17-JUN-2003
DEFINITION  Homo sapiens genomic DNA, chromosome 11 clone:CTC-1337H24, complete
            sequence.
ACCESSION   AP003419
VERSION     AP003419.3  GI:31790661
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Homo sapiens genomic DNA
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 168565)
  AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     On Jun 16, 2003 this sequence version replaced gi:22202828.
FEATURES             Location/Qualifiers
     source          1. .168565
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q"
                     /clone="CTC-1337H24"
ORIGIN

  Query Match             96.5%;  Score 4910.4;  DB 9;  Length 168565;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 4947;  Conservative    0;  Mismatches    6;  Indels    3;  Gaps    3;
```

```
Qy        133 CAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 192
              || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49326 CAAGCTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 49267

Qy        193 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49266 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 49207

Qy        253 CATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 311
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49206 CATCCTTGCTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 49147

Qy        312 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49146 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 49087

Qy        372 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49086 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 49027

Qy        432 TGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 491
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49026 TGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 48967

Qy        492 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48966 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 48907

Qy        552 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48906 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 48847

Qy        612 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48846 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 48787

Qy        672 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48786 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 48727

Qy        732 CACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCCCCA 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48726 CACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCCCCA 48667

Qy        792 TAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCTGTT 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48666 TAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCTGTT 48607

Qy        852 CTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTCAGCCTGCCAGC 911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48606 CTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGGGCTGACTGTGCTCAGCCTGCCAGC 48547

Qy        912 TCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGGGACTGGAAGAC 971
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db      48546 TCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGGGACT-GAAGAC 48488

Qy        972 TGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGTCCCTAGGCCTG 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48487 TGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGTCCCTAGGCCTG 48428

Qy       1032 GCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTGGCACACATGCC 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48427 GCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTGGCACACATGCC 48368

Qy       1092 AAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGATAGACACTCATG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48367 AAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGATAGACACTCATG 48308

Qy       1152 CGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCCGTCTTGCCAGG 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      48307 CGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCCGTCTTGCCAGG 48248
```

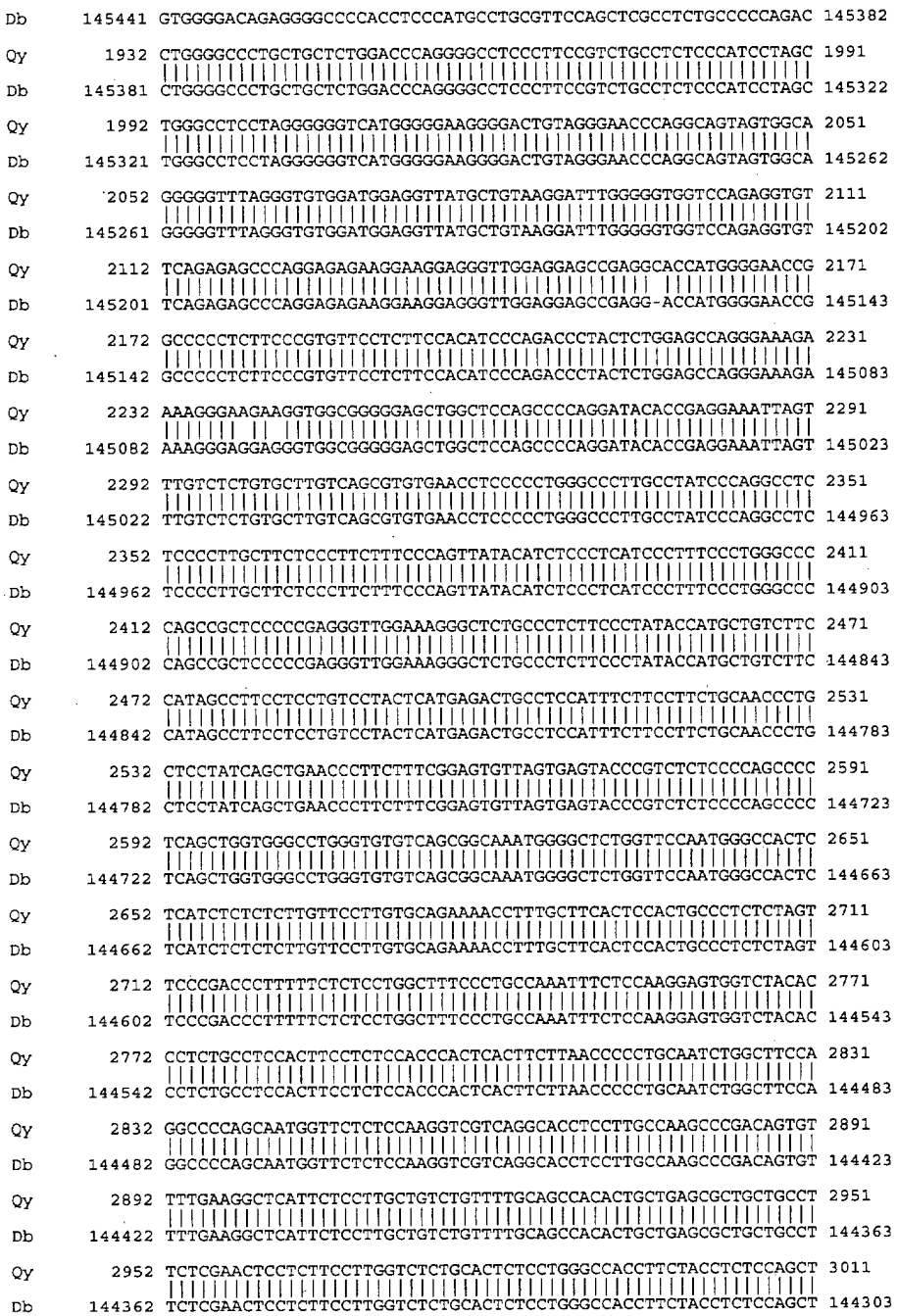
```
Qy      1212 AGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCCCCCAC 1271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48247 AGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCCCCCAC 48188

Qy      1272 CCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTGGGCCCG 1331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48187 CCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTGGGCCCG 48128

Qy      1332 GGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTAGCGT 1391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48127 GGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTAGCGT 48068

Qy      1392 GCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACC 1451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48067 GCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACC 48008

Qy      1452 CAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAAC 1511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     48007 CAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAAC 47948

Qy      1512 TCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTTGGGGAGG 1571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47947 TCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTTGGGGAGG 47888

Qy      1572 AGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGATGGGGTGA 1631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47887 AGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGATGGGGTGA 47828

Qy      1632 GGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCTTCCCTCT 1691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47827 GGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCTTCCCTCT 47768

Qy      1692 TAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCAGGCCTCG 1751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47767 TAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCAGGCCTCG 47708

Qy      1752 CACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCCCCTGGGG 1811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47707 CACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCCCCTGGGG 47648

Qy      1812 GCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCGGGGTTGG 1871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47647 GCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCGGGGTTGG 47588

Qy      1872 GTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCCCCCAGAC 1931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47587 GTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCCCCCAGAC 47528

Qy      1932 CTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCATCCTAGC 1991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47527 CTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCATCCTAGC 47468

Qy      1992 TGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGTAGTGGCA 2051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47467 TGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGTAGTGGCA 47408

Qy      2052 GGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCAGAGGTGT 2111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47407 GGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCAGAGGTGT 47348

Qy      2112 TCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACCATGGGGAACCG 2171
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     47347 TCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGG-ACCATGGGGAACCG 47289

Qy      2172 GCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAGGGAAAGA 2231
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47288 GCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAGGGAAAGA 47229

Qy      2232 AAAGGGAAGAAGGTGGCGGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGAAATTAGT 2291
             |||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47228 AAAGGGAGGAGGGTGGCGGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGAAATTAGT 47169

Qy      2292 TTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCCAGGCCTC 2351
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47168 TTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCCAGGCCTC 47109

Qy      2352 TCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCCTGGGCCC 2411
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47108 TCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCCTGGGCCC 47049

Qy      2412 CAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATGCTGTCTTC 2471
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     47048 CAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATGCTGTCTTC 46989

Qy      2472 CATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTGCAACCCTG 2531
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46988 CATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTGCAACCCTG 46929

Qy      2532 CTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCCCCAGCCCC 2591
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46928 CTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCCCCAGCCCC 46869

Qy      2592 TCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATGGGCCACTC 2651
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46868 TCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATGGGCCACTC 46809

Qy      2652 TCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCCTCTCTAGT 2711
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46808 TCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCCTCTCTAGT 46749

Qy      2712 TCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTGGTCTACAC 2771
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46748 TCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTGGTCTACAC 46689

Qy      2772 CCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCTGGCTTCCA 2831
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46688 CCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCTGGCTTCCA 46629

Qy      2832 GGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCCGACAGTGT 2891
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46628 GGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCCGACAGTGT 46569

Qy      2892 TTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCTGCTGCCT 2951
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46568 TTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCTGCTGCCT 46509

Qy      2952 TCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTCTCCAGCT 3011
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46508 TCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTCTCCAGCT 46449

Qy      3012 CCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAAGGTTTGC 3071
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46448 CCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAAGGTTTGC 46389

Qy      3072 CCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCCTCCTTTT 3131
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46388 CCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCCTCCTTTT 46329

Qy      3132 TCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCTGCATACA 3191
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46328 TCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCTGCATACA 46269

Qy      3192 AATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGCATGGAGA 3251
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46268 AATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGCATGGAGA 46209

Qy      3252 CAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTCCTTGGTG 3311
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46208 CAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTCCTTGGTG 46149

Qy      3312 GTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAGCTGAACT 3371
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46148 GTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAGCTGAACT 46089

Qy      3372 ACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTC 3431
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      46088 ACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTC 46029
Qy       3432 TGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGA 3491
Db      46028 TGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGA 45969
Qy       3492 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGG 3551
Db      45968 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGG 45909
Qy       3552 CTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGC 3611
Db      45908 CTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGC 45849
Qy       3612 TGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTG 3671
Db      45848 TGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTG 45789
Qy       3672 GGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACG 3731
Db      45788 GGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACG 45729
Qy       3732 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACC 3791
Db      45728 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACC 45669
Qy       3792 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATG 3851
Db      45668 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATG 45609
Qy       3852 GCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTTTGTG 3911
Db      45608 GCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTTTGTG 45549
Qy       3912 AGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGG 3971
Db      45548 AGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGG 45489
Qy       3972 CCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCA 4031
Db      45488 CCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCA 45429
Qy       4032 AAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGA 4091
Db      45428 AAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGA 45369
Qy       4092 AGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCT 4151
Db      45368 AGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCT 45309
Qy       4152 GCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGG 4211
Db      45308 GCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGG 45249
Qy       4212 CAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCC 4271
Db      45248 CAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCC 45189
Qy       4272 ACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGG 4331
Db      45188 ACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGG 45129
Qy       4332 AAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGG 4391
Db      45128 AAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGG 45069
Qy       4392 TCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGA 4451
Db      45068 TCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGA 45009
Qy       4452 CAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGA 4511
Db      45008 CAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGA 44949
```

```
Qy       4512 GCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTT 4571
Db      44948 GCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTT 44889
Qy       4572 TCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATT 4631
Db      44888 TCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATT 44829
Qy       4632 TCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCC 4691
Db      44828 TCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCC 44769
Qy       4692 TGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTG 4751
Db      44768 TGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTG 44709
Qy       4752 GAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAA 4811
Db      44708 GAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAA 44649
Qy       4812 AATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCTCATCAGA 4871
Db      44648 AATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCTCATCAGA 44589
Qy       4872 CTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAGAGCTCTG 4931
Db      44588 CTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAGAGCTCTG 44529
Qy       4932 GGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTACCCACTG 4991
Db      44528 GGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTACCCACTG 44469
Qy       4992 CTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTCCACCCCA 5051
Db      44468 CTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTCCACCCCA 44409
Qy       5052 GGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
Db      44408 GGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 44373
```

```
RESULT 6
AC005849/c
LOCUS       AC005849              169144 bp    DNA     linear   HTG 30-AUG-2001
DEFINITION  Homo sapiens chromosome 11 clone CIT-HSP-1337H24, 9 unordered
            pieces.
ACCESSION   AC005849
VERSION     AC005849.1  GI:3779013
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169144)
  AUTHORS   Smith,D.R.
  TITLE     Sequencing of Human Chromosome 10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169144)
  AUTHORS   Smith,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1998) Genome Therapeutics Corporation, 100 Beaver
            Street, Waltham, MA 02154, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1      506: contig of 506 bp in length
            *                   gap of unknown length
```

```
            *      507     3033: contig of 2527 bp in length
            *                   gap of unknown length
            *     3034    12956: contig of 9923 bp in length
            *                   gap of unknown length
            *    12957    19521: contig of 6565 bp in length
            *                   gap of unknown length
            *    19522    31636: contig of 12115 bp in length
            *                   gap of unknown length
            *    31637    48670: contig of 17034 bp in length
            *                   gap of unknown length
            *    48671    61610: contig of 12940 bp in length
            *                   gap of unknown length
            *    61611    97329: contig of 35719 bp in length
            *                   gap of unknown length
            *    97330   169144: contig of 71815 bp in length.
FEATURES             Location/Qualifiers
     source          1. .169144
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /clone="CIT-HSP-1337H24"
ORIGIN

  Query Match           96.5%;  Score 4908.8;  DB 2;  Length 169144;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 4946;  Conservative    0;  Mismatches    7;  Indels    3;  Gaps    3;

Qy        133 CAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 192
              || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147180 CAAGCTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 147121

Qy        193 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147120 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 147061

Qy        253 CATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 311
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147060 CATCCTTGCTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 147001

Qy        312 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147000 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 146941

Qy        372 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146940 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 146881

Qy        432 TGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 491
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146880 TGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 146821

Qy        492 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146820 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 146761

Qy        552 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146760 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 146701

Qy        612 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146700 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 146641

Qy        672 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146640 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 146581

Qy        732 CACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCCCCA 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146580 CACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCCCCA 146521

Qy        792 TAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCTGTT 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146520 TAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCTGTT 146461

Qy        852 CTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTCAGCCTGCCAGC 911
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db     146460 CTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGGGCTGACTGTGCTCAGCCTGCCAGC 146401

Qy        912 TCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGGGACTGGAAGAC 971
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db     146400 TCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGGGACT-GAAGAC 146342

Qy        972 TGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGTCCCTAGGCCTG 1031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146341 TGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGTCCCTAGGCCTG 146282

Qy       1032 GCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTGGCACACATGCC 1091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146281 GCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTGGCACACATGCC 146222

Qy       1092 AAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGATAGACACTCATG 1151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146221 AAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGATAGACACTCATG 146162

Qy       1152 CGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCCGTCTTGCCAGG 1211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146161 CGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCCGTCTTGCCAGG 146102

Qy       1212 AGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCCCCCAC 1271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146101 AGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCCCCCAC 146042

Qy       1272 CCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTGGGCCCG 1331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146041 CCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTGGGCCCG 145982

Qy       1332 GGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTAGCGT 1391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145981 GGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATGTTAGCGT 145922

Qy       1392 GCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACC 1451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145921 GCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACC 145862

Qy       1452 CAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAAC 1511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145861 CAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAGCACCAAC 145802

Qy       1512 TCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTTGGGGAGG 1571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145801 TCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTTGGGGAGG 145742

Qy       1572 AGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGATGGGGTGA 1631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145741 AGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGATGGGGTGA 145682

Qy       1632 GGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCTTCCCTCT 1691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145681 GGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCTTCCCTCT 145622

Qy       1692 TAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCAGGCCTCG 1751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145621 TAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCAGGCCTCG 145562

Qy       1752 CACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCCCCTGGGG 1811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145561 CACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCCCCTGGGG 145502

Qy       1812 GCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCGGGGTTGG 1871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     145501 GCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCGGGGTTGG 145442

Qy       1872 GTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCCCCCAGAC 1931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

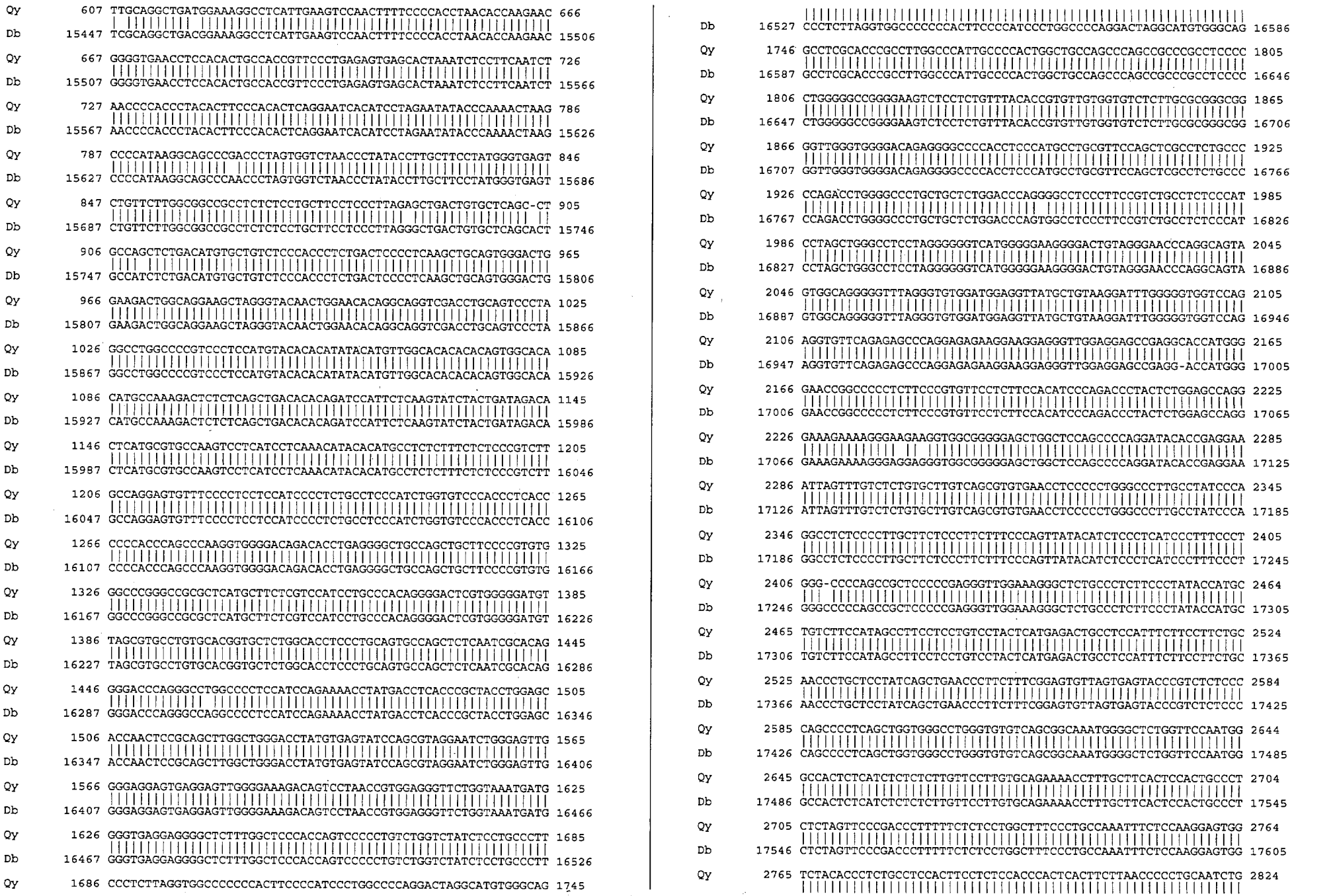

Db 145441 GTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCCCCCAGAC 145382
Qy 1932 CTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCATCCTAGC 1991
Db 145381 CTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCATCCTAGC 145322
Qy 1992 TGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGTAGTGGCA 2051
Db 145321 TGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGTAGTGGCA 145262
Qy 2052 GGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCAGAGGTGT 2111
Db 145261 GGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCAGAGGTGT 145202
Qy 2112 TCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACCATGGGGAACCG 2171
Db 145201 TCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGG-ACCATGGGGAACCG 145143
Qy 2172 GCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAGGGAAAGA 2231
Db 145142 GCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAGGGAAAGA 145083
Qy 2232 AAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGAAATTAGT 2291
Db 145082 AAAGGGAGGAGGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGAAATTAGT 145023
Qy 2292 TTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCCAGGCCTC 2351
Db 145022 TTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCCAGGCCTC 144963
Qy 2352 TCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCCTGGGCCC 2411
Db 144962 TCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCCTGGGCCC 144903
Qy 2412 CAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATGCTGTCTTC 2471
Db 144902 CAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATGCTGTCTTC 144843
Qy 2472 CATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTGCAACCCTG 2531
Db 144842 CATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTGCAACCCTG 144783
Qy 2532 CTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCCCCAGCCCC 2591
Db 144782 CTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCCCCAGCCCC 144723
Qy 2592 TCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATGGGCCACTC 2651
Db 144722 TCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATGGGCCACTC 144663
Qy 2652 TCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCCTCTCTAGT 2711
Db 144662 TCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCCTCTCTAGT 144603
Qy 2712 TCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTGGTCTACAC 2771
Db 144602 TCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTGGTCTACAC 144543
Qy 2772 CCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCTGGCTTCCA 2831
Db 144542 CCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCTGGCTTCCA 144483
Qy 2832 GGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCCGACAGTGT 2891
Db 144482 GGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCCGACAGTGT 144423
Qy 2892 TTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCTGCTGCCT 2951
Db 144422 TTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCTGCTGCCT 144363
Qy 2952 TCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTCTCCAGCT 3011
Db 144362 TCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTCTCCAGCT 144303
Qy 3012 CCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAAGGTTTGC 3071
Db 144302 CCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAAGGTTTGC 144243
Qy 3072 CCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCCTCCTTTT 3131
Db 144242 CCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCCTCCTTTT 144183
Qy 3132 TCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCTGCATACA 3191
Db 144182 TCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCTGCATACA 144123
Qy 3192 AATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGCATGGAGA 3251
Db 144122 AATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGCATGGAGA 144063
Qy 3252 CAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTCCTTGGTG 3311
Db 144062 CAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTCCTTGGTG 144003
Qy 3312 GTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAGCTGAACT 3371
Db 144002 GTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAGCTGAACT 143943
Qy 3372 ACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTC 3431
Db 143942 ACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAAAGACTC 143883
Qy 3432 TGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGA 3491
Db 143882 TGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGA 143823
Qy 3492 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGG 3551
Db 143822 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGG 143763
Qy 3552 CTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGC 3611
Db 143762 CTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGC 143703
Qy 3612 TGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTG 3671
Db 143702 TGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTG 143643
Qy 3672 GGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACG 3731
Db 143642 GGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACG 143583
Qy 3732 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACC 3791
Db 143582 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACC 143523
Qy 3792 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATG 3851
Db 143522 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATG 143463
Qy 3852 GCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTTTGTG 3911
Db 143462 GCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTTTGTG 143403
Qy 3912 AGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGG 3971
Db 143402 AGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGG 143343
Qy 3972 CCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCA 4031
Db 143342 CCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCA 143283
Qy 4032 AAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGA 4091
Db 143282 AAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGA 143223

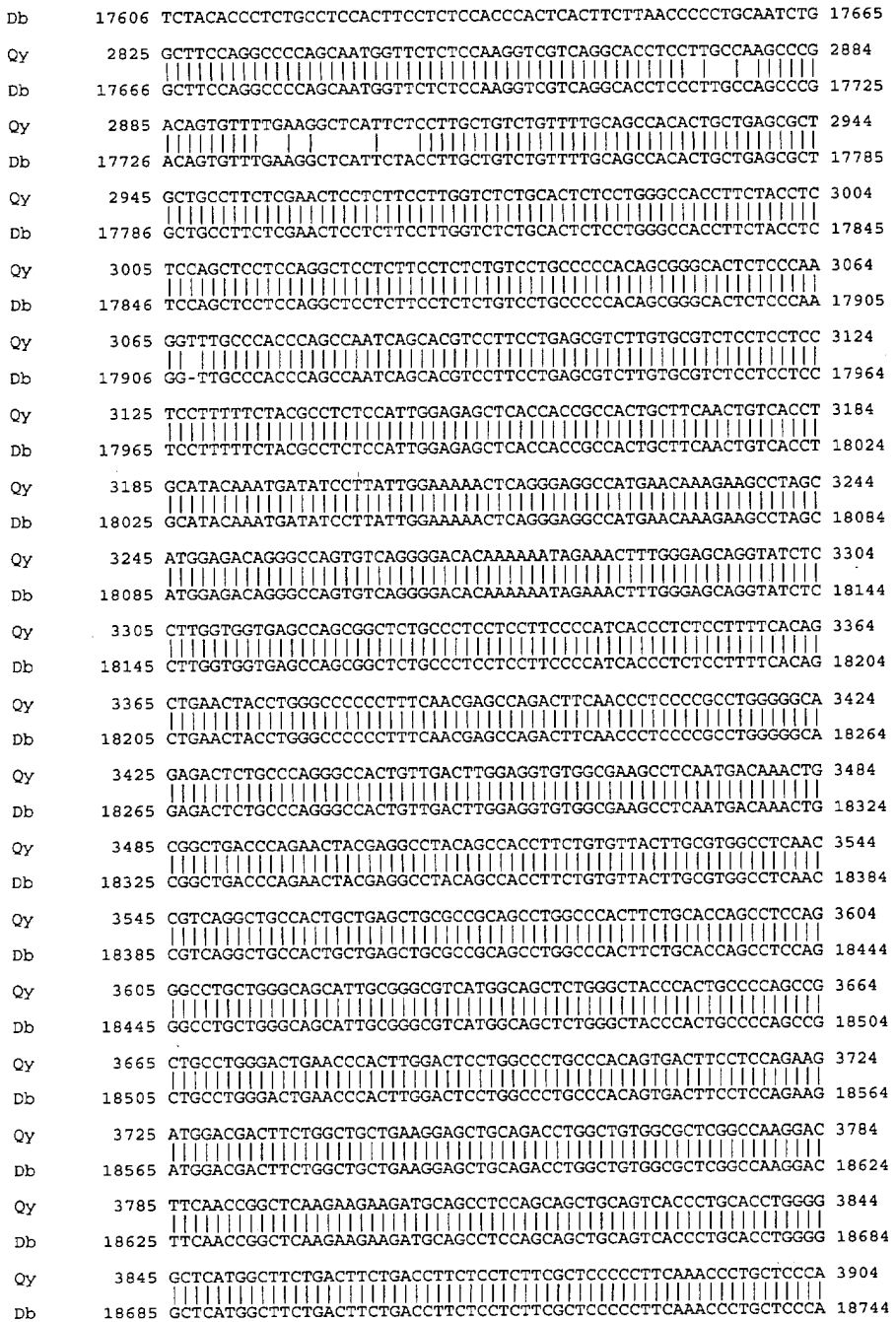
```
Qy       4092 AGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCT 4151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143222 AGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCT 143163

Qy       4152 GCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGG 4211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143162 GCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGG 143103

Qy       4212 CAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCC 4271
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143102 CAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCC 143043

Qy       4272 ACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGG 4331
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143042 ACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGG 142983

Qy       4332 AAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGG 4391
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142982 AAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGG 142923

Qy       4392 TCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGA 4451
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142922 TCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGA 142863

Qy       4452 CAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGA 4511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142862 CAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGA 142803

Qy       4512 GCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTT 4571
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142802 GCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTT 142743

Qy       4572 TCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATT 4631
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142742 TCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATT 142683

Qy       4632 TCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCC 4691
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142682 TCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCC 142623

Qy       4692 TGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTG 4751
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142622 TGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTG 142563

Qy       4752 GAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAA 4811
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142562 GAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAA 142503

Qy       4812 AATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCTCATCAGA 4871
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142502 AATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCTCATCAGA 142443

Qy       4872 CTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAGAGCTCTG 4931
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142442 CTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAGAGCTCTG 142383

Qy       4932 GGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTACCCACTG 4991
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142382 GGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTACCCACTG 142323

Qy       4992 CTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTCCACCCCA 5051
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142322 CTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTCCACCCCA 142263

Qy       5052 GGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||
Db     142262 GGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 142227
```

RESULT 7
AP002437
LOCUS AP002437 135116 bp DNA linear HTG 06-JUN-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-678D20 map 11q13, WORKING
DRAFT SEQUENCE, 41 unordered pieces.
ACCESSION AP002437
VERSION AP002437.1 GI:8307741
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135116)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 135,116 genomic DNA of 11q13
JOURNAL Published Only in DataBase (2000)
REFERENCE 2 (bases 1 to 135116)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923, Fax:81-42-778-9924)
COMMENT ---------- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
---------- Project Information
Center project name: HumDraft11
Center clone name: RP11-678D20
---------- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 109313 bases at least Q40
Consensus quality: 121923 bases at least Q30
Consensus quality: 127981 bases at least Q20
Insert size: 131116; sum-of-contigs
Quality coverage: 4.18x in Q20 bases; sum-of-contigs
----------.
NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

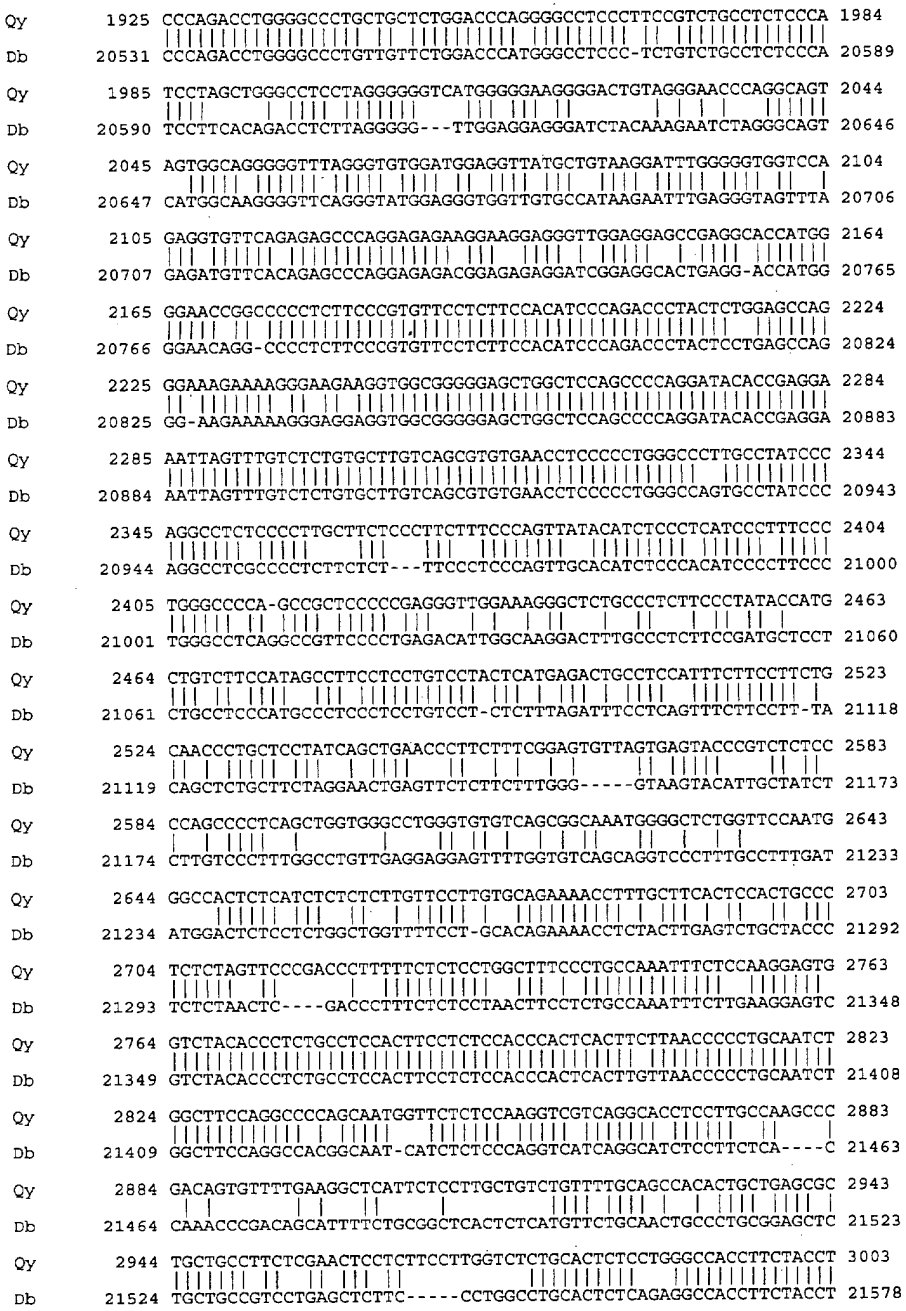
1	7749	contig of	7749	bp in length
7850	15108	contig of	7259	bp in length
15209	23728	contig of	8520	bp in length
23829	29302	contig of	5474	bp in length
29403	36171	contig of	6769	bp in length
36272	42123	contig of	5852	bp in length
42224	48172	contig of	5949	bp in length
48273	52700	contig of	4428	bp in length
52801	57587	contig of	4787	bp in length
57688	61624	contig of	3937	bp in length
61725	67756	contig of	6032	bp in length
67857	71834	contig of	3978	bp in length
71935	76800	contig of	4866	bp in length
76901	81069	contig of	4169	bp in length
81170	81932	contig of	763	bp in length
82033	86213	contig of	4181	bp in length
86314	89537	contig of	3224	bp in length
89638	92372	contig of	2735	bp in length
92473	94979	contig of	2507	bp in length
95080	96272	contig of	1193	bp in length
96373	98109	contig of	1737	bp in length
98210	101030	contig of	2821	bp in length
101131	103760	contig of	2630	bp in length
103861	106330	contig of	2470	bp in length
106431	108601	contig of	2171	bp in length

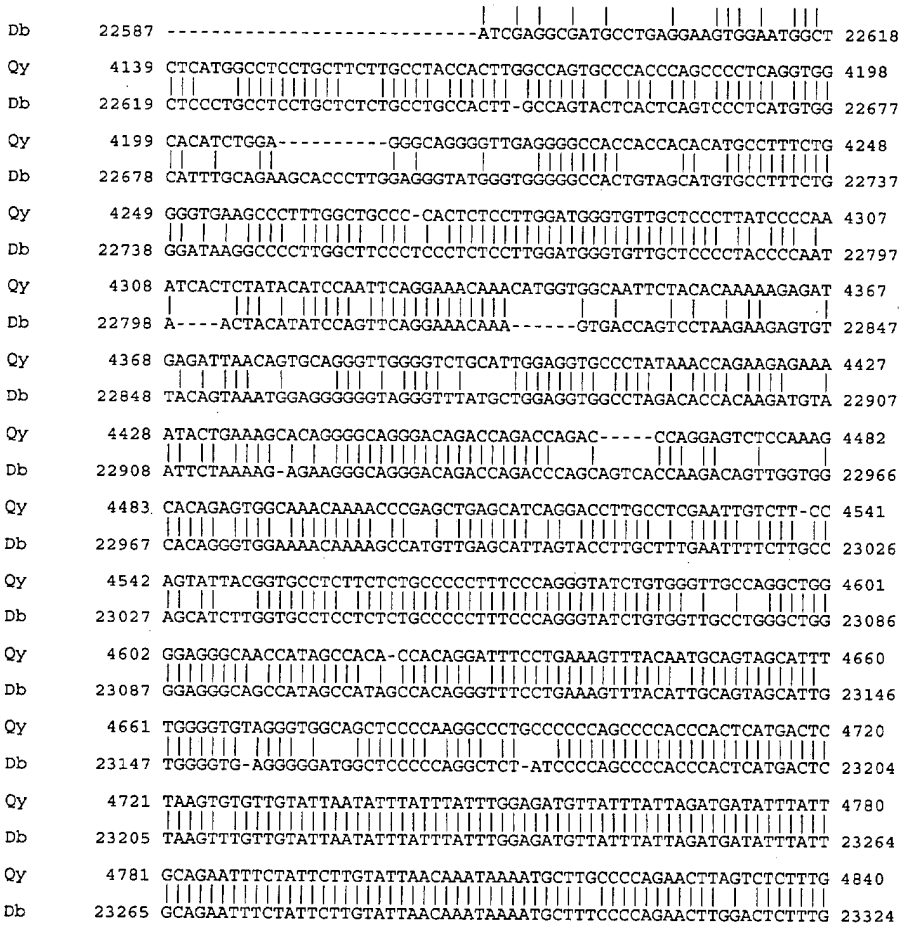
```
    108702    110440 contig of   1739 bp in length
    110541    112280 contig of   1740 bp in length
    112381    114885 contig of   2505 bp in length
    114986    117018 contig of   2033 bp in length
    117119    118839 contig of   1721 bp in length
    118940    120968 contig of   2029 bp in length
    121069    122947 contig of   1879 bp in length
    123048    124231 contig of   1184 bp in length
    124332    125727 contig of   1396 bp in length
    125828    127192 contig of   1365 bp in length
    127293    129019 contig of   1727 bp in length
    129120    130181 contig of   1062 bp in length
    130282    131483 contig of   1202 bp in length
    131584    132595 contig of   1012 bp in length
    132696    133886 contig of   1191 bp in length
    133987    135116 contig of   1130 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 41 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*         1      7749: contig of 7749 bp in length
*      7750      7849: gap of 100 bp
*      7850     15108: contig of 7259 bp in length
*     15109     15208: gap of 100 bp
*     15209     23728: contig of 8520 bp in length
*     23729     23828: gap of 100 bp
*     23829     29302: contig of 5474 bp in length
*     29303     29402: gap of 100 bp
*     29403     36171: contig of 6769 bp in length
*     36172     36271: gap of 100 bp
*     36272     42123: contig of 5852 bp in length
*     42124     42223: gap of 100 bp
*     42224     48172: contig of 5949 bp in length
*     48173     48272: gap of 100 bp
*     48273     52700: contig of 4428 bp in length
*     52701     52800: gap of 100 bp
*     52801     57587: contig of 4787 bp in length
*     57588     57687: gap of 100 bp
*     57688     61624: contig of 3937 bp in length
*     61625     61724: gap of 100 bp
*     61725     67756: contig of 6032 bp in length
*     67757     67856: gap of 100 bp
*     67857     71834: contig of 3978 bp in length
*     71835     71934: gap of 100 bp
*     71935     76800: contig of 4866 bp in length
*     76801     76900: gap of 100 bp
*     76901     81069: contig of 4169 bp in length
*     81070     81169: gap of 100 bp
*     81170     81932: contig of 763 bp in length
*     81933     82032: gap of 100 bp
*     82033     86213: contig of 4181 bp in length
*     86214     86313: gap of 100 bp
*     86314     89537: contig of 3224 bp in length
*     89538     89637: gap of 100 bp
*     89638     92372: contig of 2735 bp in length
*     92373     92472: gap of 100 bp
*     92473     94979: contig of 2507 bp in length
*     94980     95079: gap of 100 bp
*     95080     96272: contig of 1193 bp in length
*     96273     96372: gap of 100 bp
*     96373     98109: contig of 1737 bp in length
*     98110     98209: gap of 100 bp
*     98210    101030: contig of 2821 bp in length
*    101031    101130: gap of 100 bp
*    101131    103760: contig of 2630 bp in length
*    103761    103860: gap of 100 bp
*    103861    106330: contig of 2470 bp in length
*    106331    106430: gap of 100 bp
*    106431    108601: contig of 2171 bp in length
*    108602    108701: gap of 100 bp
*    108702    110440: contig of 1739 bp in length
*    110441    110540: gap of 100 bp
*    110541    112280: contig of 1740 bp in length
*    112281    112380: gap of 100 bp
*    112381    114885: contig of 2505 bp in length
*    114886    114985: gap of 100 bp
*    114986    117018: contig of 2033 bp in length
*    117019    117118: gap of 100 bp
*    117119    118839: contig of 1721 bp in length
*    118840    118939: gap of 100 bp
*    118940    120968: contig of 2029 bp in length
*    120969    121068: gap of 100 bp
*    121069    122947: contig of 1879 bp in length
*    122948    123047: gap of 100 bp
*    123048    124231: contig of 1184 bp in length
*    124232    124331: gap of 100 bp
*    124332    125727: contig of 1396 bp in length
*    125728    125827: gap of 100 bp
*    125828    127192: contig of 1365 bp in length
*    127193    127292: gap of 100 bp
*    127293    129019: contig of 1727 bp in length
*    129020    129119: gap of 100 bp
*    129120    130181: contig of 1062 bp in length
*    130182    130281: gap of 100 bp
*    130282    131483: contig of 1202 bp in length
*    131484    131583: gap of 100 bp
*    131584    132595: contig of 1012 bp in length
*    132596    132695: gap of 100 bp
*    132696    133886: contig of 1191 bp in length
*    133887    133986: gap of 100 bp
*    133987    135116: contig of 1130 bp in length.
FEATURES             Location/Qualifiers
     source          1. .135116
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13"
                     /clone="RP11-678D20"
     misc_feature    1. .7749
                     /note="assembly_fragment"
     misc_feature    7850. .15108
                     /note="assembly_fragment"
     misc_feature    15209. .23728
                     /note="assembly_fragment"
     misc_feature    23829. .29302
                     /note="assembly_fragment clone_end:SP6 vector_side:right"
     misc_feature    29403. .36171
                     /note="assembly_fragment"
     misc_feature    36272. .42123
                     /note="assembly_fragment"

  Query Match             90.5%;  Score 4605;  DB 2;  Length 135116;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 4682;  Conservative    0;  Mismatches   35;  Indels    6;  Gaps    5;

Qy        367 TATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCT 426
              ||||||||||||||||||||||| ||| || ||||||||||||||||||||||||||||
Db      15209 TATATCTACCCTAGCCAAACCCTAGCTGACTATGCCAGTCCAAAGTGTCCACTGAATCCT 15268

Qy        427 CACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCC 486
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      15269 CACCTTGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCC 15328

Qy        487 TTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAG 546
              |||||||||||| |||||||||||||||||||||||||  ||     ||||||||||||
Db      15329 TTACAGCACCCAAGCCTGGCCTCTGGACTCCTGGTATC--TGGGATGTCAAACTCTGCAG 15386

Qy        547 TGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCC 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15387 TGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCC 15446
```




Db 17606 TCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCTG 17665
Qy 2825 GCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCCG 2884
Db 17666 GCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCCTTGCCAGCCCG 17725
Qy 2885 ACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCT 2944
Db 17726 ACAGTGTTTGAAGGCTCATTCTACCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGCT 17785
Qy 2945 GCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTC 3004
Db 17786 GCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCTC 17845
Qy 3005 TCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAA 3064
Db 17846 TCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCAA 17905
Qy 3065 GGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCC 3124
Db 17906 GG-TTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTCC 17964
Qy 3125 TCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCT 3184
Db 17965 TCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACCT 18024
Qy 3185 GCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGC 3244
Db 18025 GCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAGC 18084
Qy 3245 ATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTC 3304
Db 18085 ATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCTC 18144
Qy 3305 CTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAG 3364
Db 18145 CTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAG 18204
Qy 3365 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 3424
Db 18205 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 18264
Qy 3425 GAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 3484
Db 18265 GAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 18324
Qy 3485 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 3544
Db 18325 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 18384
Qy 3545 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 3604
Db 18385 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 18444
Qy 3605 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 3664
Db 18445 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 18504
Qy 3665 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 3724
Db 18505 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 18564
Qy 3725 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 3784
Db 18565 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 18624
Qy 3785 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 3844
Db 18625 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 18684
Qy 3845 GCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCA 3904
Db 18685 GCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCA 18744
Qy 3905 CTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAG 3964
Db 18745 CTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAG 18804
Qy 3965 CCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCA 4024
Db 18805 CCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCA 18864
Qy 4025 CCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTC 4084
Db 18865 CCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTC 18924
Qy 4085 TACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATG 4144
Db 18925 TACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATG 18984
Qy 4145 GCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATC 4204
Db 18985 GCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATC 19044
Qy 4205 TGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGG 4264
Db 19045 TGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGG 19104
Qy 4265 CTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCA 4324
Db 19105 CTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCA 19164
Qy 4325 ATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGG 4384
Db 19165 ATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGG 19224
Qy 4385 GTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGG 4444
Db 19225 GTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGG 19284
Qy 4445 GCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCC 4504
Db 19285 GCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCC 19344
Qy 4505 GAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTG 4564
Db 19345 GAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTG 19404
Qy 4565 CCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCA 4624
Db 19405 CCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCA 19464
Qy 4625 CAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCC 4684
Db 19465 CAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCC 19524
Qy 4685 AAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTAT 4744
Db 19525 AAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTAT 19584
Qy 4745 TTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAA 4804
Db 19585 TTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAA 19644
Qy 4805 CAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCT 4864
Db 19645 CAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGGTGCT 19704
Qy 4865 CATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAG 4924
Db 19705 CATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGCACAG 19764
Qy 4925 AGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTA 4984
Db 19765 AGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGACTTA 19824

```
Qy       4985 CCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTC 5044
Db      19825 CCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTCTGTC 19884

Qy       5045 CACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
Db      19885 CACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 19927

RESULT 8
AC140073
LOCUS       AC140073            246093 bp    DNA     linear   HTG 07-OCT-2003
DEFINITION  Mus musculus chromosome UNK clone RP23-282N20, WORKING DRAFT
            SEQUENCE, 5 unordered pieces.
ACCESSION   AC140073
VERSION     AC140073.3  GI:37537681
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 246093)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 246093)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 246093)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Oct 7, 2003 this sequence version replaced gi:28475471.

            ----------------- Genome Center ---------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information ------------------
            Center project name: M_BA0282N20
            -----------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; %
            Sequencing vector: plasmid; %
            Chemistry: Dye-primer ET; % of reads
            Chemistry: Dye-terminator Big Dye; % of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 241933 bases at least Q40
            Consensus quality: 242675 bases at least Q30
            Consensus quality: 243192 bases at least Q20
            Insert size: 182000; agarose-fp
            Insert size: 247285; sum-of-contigs
            Quality coverage: 11.88 in Q20 bases; agarose-fp
            Quality coverage: 10.46 in Q20 bases; sum-of-contigs
            -----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    4448: contig of 4448 bp in length
            *     4449    4548: gap of unknown length
            *     4549   38123: contig of 33575 bp in length
            *    38124   38223: gap of unknown length
            *    38224   65661: contig of 27438 bp in length
            *    65662   65761: gap of unknown length
            *    65762  128916: contig of 63155 bp in length
            *   128917  129016: gap of unknown length
            *   129017  246093: contig of 117077 bp in length.
FEATURES             Location/Qualifiers
     source          1..246093
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="UNK"
                     /clone="RP23-282N20"
     misc_feature    1..4448
                     /note="assembly_name:Contig21"
     misc_feature    4549..38123
                     /note="assembly_name:Contig22"
     misc_feature    38224..65661
                     /note="assembly_name:Contig23"
     misc_feature    65762..128916
                     /note="assembly_name:Contig24"
     misc_feature    129017..246093
                     /note="assembly_name:Contig25"
ORIGIN

  Query Match             33.6%;  Score 1709.4;  DB 2;  Length 246093;
  Best Local Similarity   74.5%;  Pred. No. 0;
  Matches 2733;  Conservative    0;  Mismatches  742;  Indels  193;  Gaps   39;

Qy       1207 CCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCC 1266
Db      19832 CCAGAAGTCTTTCCCCTCCTCCACCCCCTCCATTTCCCATCTGGTCTCTCTCCCTCACCC 19891

Qy       1267 CCC-ACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTG 1325
Db      19892 CCCTTCCTAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGTTTCCCTATTTG 19951

Qy       1326 GGCCCGGG-CCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATG 1384
Db      19952 GGCCTGGGCCCACCCTCATACTTCTTGCCCATCCTATTCACAGGGGACTCGTGGGGGATG 20011

Qy       1385 TTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACA 1444
Db      20012 TTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGCACA 20071

Qy       1445 GGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAG 1504
Db      20072 GGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAG 20131

Qy       1505 CACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTT 1564
Db      20132 CATCAACTCCGCAGCTTAGCTGGGACCTACGTGAGTATCCCCTTTGGCAA---------- 20181

Qy       1565 GGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGAT 1624
Db      20182 --AGAGGAGGCAGAGGCTTTGGGGAAGAGAGAGCATGGGGCAGGGGTCCTAGTGAATGAT 20239

Qy       1625 GGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCT 1684
Db      20240 GGGGTGATGAGAGGCCCTTTGGGTCCACATAGCTTCTCCATCTGGCCTATCTTCTGCCCT 20299

Qy       1685 TCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCA 1744
Db      20300 TCCCTCTCAGGTGCCCCCC-------CCCAATCCTGGCCCCAGGAGTAGGCATGTGGGCA 20352

Qy       1745 GGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCC 1804
Db      20353 GGCCTCGAACCCGCCTTGGCCCATTGCCCCATTGGCTGCCAGCCCAGCTGCCTGCCTCCC 20412

Qy       1805 CCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCG 1864
Db      20413 CCTGGGGGCTGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCAAGGGTG 20472

Qy       1865 GGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCC 1924
Db      20473 GGGCTGGGTGGGGATGGAGGGGCCCCACCTTCTATGCTTGTGTTCCAGCTCGC--CTACC 20530
```

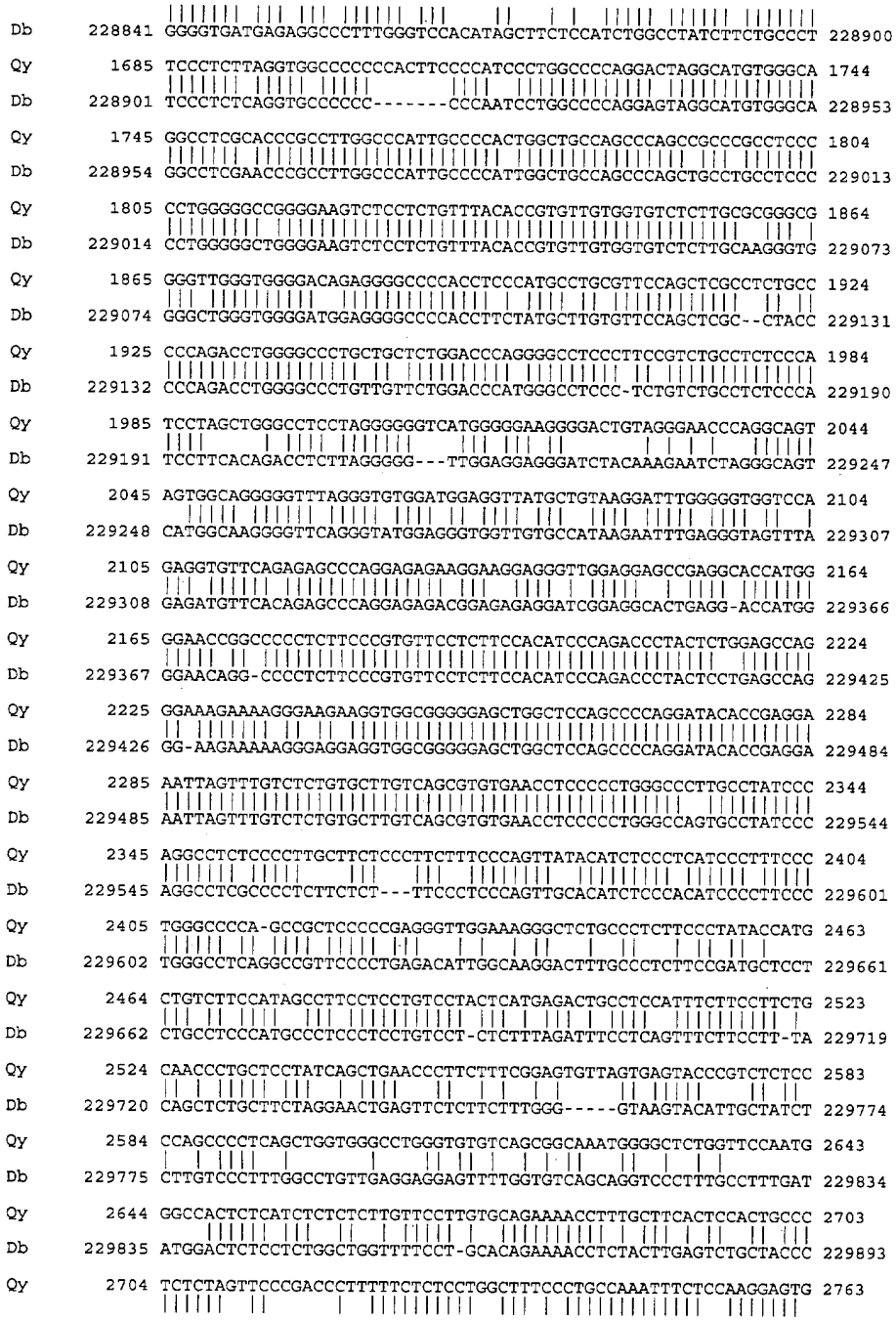

Qy 1925 CCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCA 1984
Db 20531 CCCAGACCTGGGGCCCTGTTGTTCTGGACCCATGGGCCTCCC-TCTGTCTGCCTCTCCCA 20589
Qy 1985 TCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGT 2044
Db 20590 TCCTTCACAGACCTCTTAGGGGG---TTGGAGGAGGGATCTACAAAGAATCTAGGGCAGT 20646
Qy 2045 AGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCA 2104
Db 20647 CATGGCAAGGGGTTCAGGGTATGGAGGGTGGTTGTGCCATAAGAATTTGAGGGTAGTTTA 20706
Qy 2105 GAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACCATGG 2164
Db 20707 GAGATGTTCACAGAGCCCAGGAGAGACGGAGAGAGGATCGGAGGCACTGAGG-ACCATGG 20765
Qy 2165 GGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAG 2224
Db 20766 GGAACAGG-CCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCCTGAGCCAG 20824
Qy 2225 GGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGA 2284
Db 20825 GG-AAGAAAAAGGGAGGAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGA 20883
Qy 2285 AATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCC 2344
Db 20884 AATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCAGTGCCTATCCC 20943
Qy 2345 AGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCC 2404
Db 20944 AGGCCTCGCCCCTCTTCTCT---TTCCCTCCCAGTTGCACATCTCCCACATCCCCTTCCC 21000
Qy 2405 TGGGCCCCA-GCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATG 2463
Db 21001 TGGGCCTCAGGCCGTTCCCCTGAGACATTGGCAAGGACTTTGCCCTCTTCCGATGCTCCT 21060
Qy 2464 CTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTG 2523
Db 21061 CTGCCTCCCATGCCCTCCCTCCTGTCCT-CTCTTTAGATTTCCTCAGTTTCTTCCTT-TA 21118
Qy 2524 CAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCC 2583
Db 21119 CAGCTCTGCTTCTAGGAACTGAGTTCTCTTCTTTGGG-----GTAAGTACATTGCTATCT 21173
Qy 2584 CCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATG 2643
Db 21174 CTTGTCCCTTTGGCCTGTTGAGGAGGAGTTTTGGTGTCAGCAGGTCCCTTTGCCTTTGAT 21233
Qy 2644 GGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCC 2703
Db 21234 ATGGACTCTCCTCTGGCTGGTTTTCCT-GCACAGAAAACCTCTACTTGAGTCTGCTACCC 21292
Qy 2704 TCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTG 2763
Db 21293 TCTCTAACTC----GACCCTTTCTCTCCTAACTTCCTCTGCCAAATTTCTTGAAGGAGTC 21348
Qy 2764 GTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCT 2823
Db 21349 GTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTGTTAACCCCCTGCAATCT 21408
Qy 2824 GGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCC 2883
Db 21409 GGCTTCCAGGCCACGGCAAT-CATCTCTCCCAGGTCATCAGGCATCTCCTTCTCA----C 21463
Qy 2884 GACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGC 2943
Db 21464 CAAACCCGACAGCATTTTCTGCGGCTCACTCTCATGTTCTGCAACTGCCCTGCGGAGCTC 21523
Qy 2944 TGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCT 3003
Db 21524 TGCTGCCGTCCTGAGCTCTTC-----CCTGGCCTGCACTCTCAGAGGCCACCTTCTACCT 21578
Qy 3004 CTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCA 3063
Db 21579 CTCCAGCCTGGTCTCCTGGCTCCTCCG-GCTCCTCTTCTTCTCTGCCCTGCCCTGTGGTG 21637
Qy 3064 AGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTC 3123
Db 21638 GGCATTCTCCACCCAGACAGTCAGCACGTCCCACCCGAGTGGCTCCTTCGTGGTTTTCTC 21697
Qy 3124 CTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACC 3183
Db 21698 ------TCTCTGCACTTCTCCATAGGAGAACTCATCACC-CCACTGCTTCACCTGTCTCC 21750
Qy 3184 TGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAG 3243
Db 21751 TCCATGCAAATGATACACTTGCTGGAAAACCT---------------------------- 21782
Qy 3244 CATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCT 3303
Db 21783 ------------CCCATGTCAGGGGTCAAGAGCGATAGAAGCTTTGGGAGGAGGTATCT 21829
Qy 3304 CCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACA 3363
Db 21830 CTCTGCTGGCAAGCTGGCAGCTCTGCTCTGCTCCTTCCCCATTGCCCTCTCCTTTTCACA 21889
Qy 3364 GCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGC 3423
Db 21890 GCTGAACTACCTGGGGCCCCCTTTCAACGAGCCTGACTTCAATCCTCCTCGACTGGGGGC 21949
Qy 3424 AGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACT 3483
Db 21950 AGAAACTCTGCCCAGGGCCACGGTCAACTTGGAAGTGTGGCGAAGCCTCAATGACAGGCT 22009
Qy 3484 GCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAA 3543
Db 22010 GCGGCTGACCCAGAACTATGAGGCGTACAGTCACCTCCTGTGTTACTTGCGTGGCCTCAA 22069
Qy 3544 CCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCA 3603
Db 22070 CCGTCAGGCTGCCACAGCTGAACTCCGACGTAGCCTGGCCCACTTCTGTACCAGCCTCCA 22129
Qy 3604 GGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCC 3663
Db 22130 GGGCCTGCTGGGCAGCATTGCAGGTGTCATGGCGACGCTTGGCTACCCACTGCCCCAGCC 22189
Qy 3664 GCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAA 3723
Db 22190 TCTGCCAGGGACTGAGCCAGCCTGGGCCCCTGGCCCTGCCCACAGTGACTTCCTCCAGAA 22249
Qy 3724 GATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGA 3783
Db 22250 GATGGATGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTATGGCGTTCAGCCAAGGA 22309
Qy 3784 CTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGG 3843
Db 22310 CTTCAACCGGCTTAAGAAGAAGATGCAGCCTCCAGCAGCTTCAGTCACCCTGCACTTGGA 22369
Qy 3844 GGCTCATGGCTTCTGACTTCTGACCTT-----CTCCTCTTCGCTCCCCCTTCAAACCCTG 3898
Db 22370 GGCCCATGGTTTCTGACCTCTGACCCTTAACCCCCACACCTCCAGGCCCAGTCAGCTGTG 22429
Qy 3899 CTCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGC 3958
Db 22430 CTTCCATTGTGGGGTGAAAAATCCTGTATACCAACATTTGTTGAGCCAGGAACCAGACCA 22489
Qy 3959 TGTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCC 4018
Db 22490 TGGGAG-CTCTGCCCTCCTCAGCCTTGGATGGGGAGTGTAACTCAATAAGCCCTCTAGC- 22547
Qy 4019 TCCCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCC 4078
Db 22548 ---------TCAGAAAGTCCTAGGGTCTGGGGAAATCGTACAGTGGGC------------ 22586
Qy 4079 TGTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCG 4138

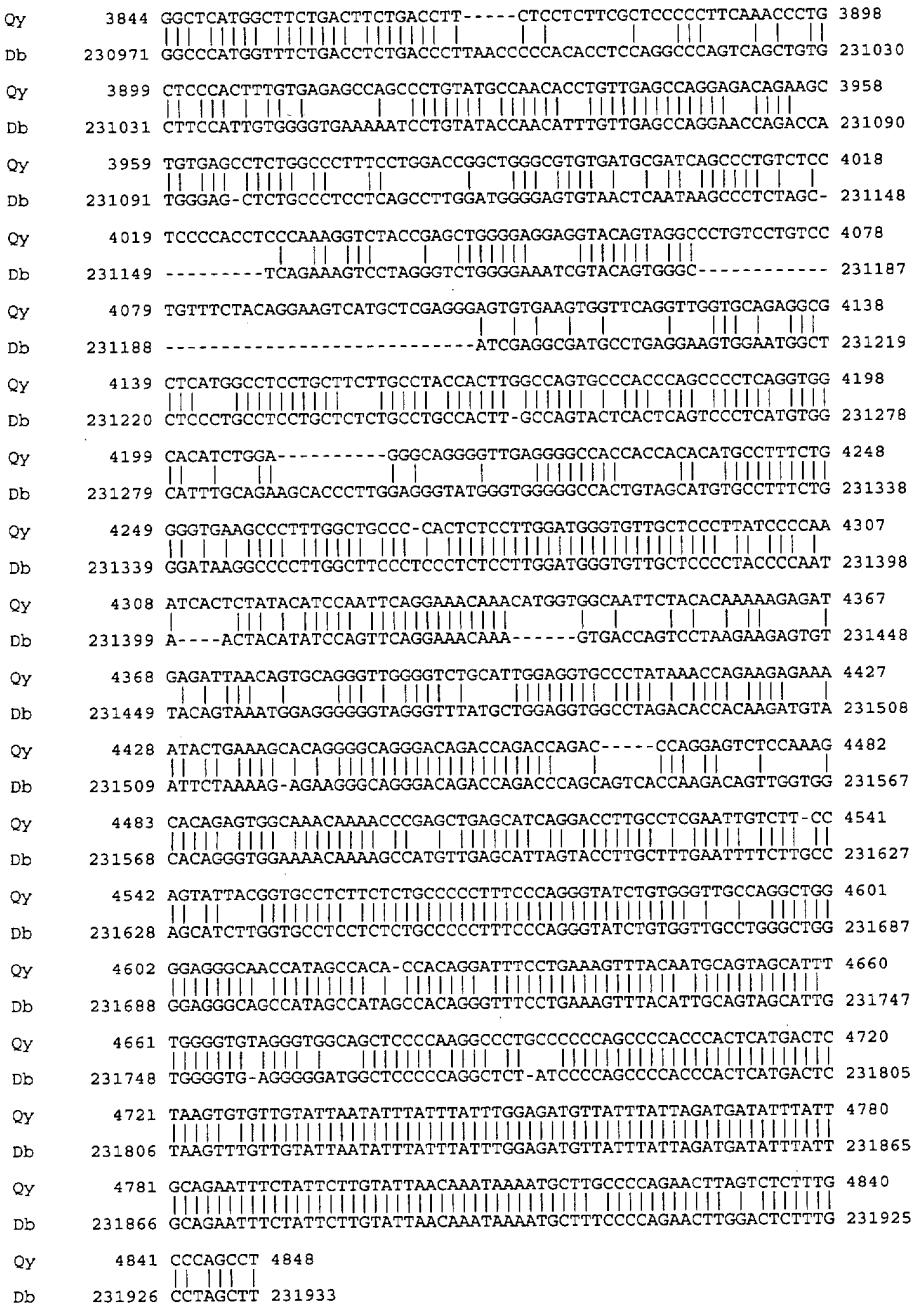


```
Db      22587 ----------------------------ATCGAGGCGATGCCTGAGGAAGTGGAATGGCT 22618

Qy       4139 CTCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGG 4198

Db      22619 CTCCCTGCCTCCTGCTCTCTGCCTGCCACTT-GCCAGTACTCACTCAGTCCCTCATGTGG 22677

Qy       4199 CACATCTGGA----------GGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTG 4248

Db      22678 CATTTGCAGAAGCACCCTTGGAGGGTATGGGTGGGGGCCACTGTAGCATGTGCCTTTCTG 22737

Qy       4249 GGGTGAAGCCCTTTGGCTGCCC-CACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAA 4307

Db      22738 GGATAAGGCCCCTTGGCTTCCCTCCCTCTCCTTGGATGGGTGTTGCTCCCCTACCCCAAT 22797

Qy       4308 ATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGAT 4367

Db      22798 A----ACTACATATCCAGTTCAGGAAACAAA------GTGACCAGTCCTAAGAAGAGTGT 22847

Qy       4368 GAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAA 4427

Db      22848 TACAGTAAATGGAGGGGGGTAGGGTTTATGCTGGAGGTGGCCTAGACACCACAAGATGTA 22907

Qy       4428 ATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGAC-----CCAGGAGTCTCCAAAG 4482

Db      22908 ATTCTAAAAG-AGAAGGGCAGGGACAGACCAGACCCAGCAGTCACCAAGACAGTTGGTGG 22966

Qy       4483 CACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTT-CC 4541

Db      22967 CACAGGGTGGAAAACAAAAGCCATGTTGAGCATTAGTACCTTGCTTTGAATTTTCTTGCC 23026

Qy       4542 AGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGG 4601

Db      23027 AGCATCTTGGTGCCTCCTCTCTGCCCCCTTTCCCAGGGTATCTGTGGTTGCCTGGGCTGG 23086

Qy       4602 GGAGGGCAACCATAGCCACA-CCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTT 4660

Db      23087 GGAGGGCAGCCATAGCCATAGCCACAGGGTTTCCTGAAAGTTTACATTGCAGTAGCATTG 23146

Qy       4661 TGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTC 4720

Db      23147 TGGGGTG-AGGGGGATGGCTCCCCCAGGCTCT-ATCCCCAGCCCCACCCACTCATGACTC 23204

Qy       4721 TAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATT 4780

Db      23205 TAAGTTTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATT 23264

Qy       4781 GCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTG 4840

Db      23265 GCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTTCCCCAGAACTTGGACTCTTTG 23324

Qy       4841 CCCAGCCT 4848

Db      23325 CCTAGCTT 23332
```

RESULT 9
AC109138
LOCUS AC109138 269155 bp DNA linear ROD 01-OCT-2003
DEFINITION Mus musculus chromosome 19, clone RP23-41B18, complete sequence.
ACCESSION AC109138
VERSION AC109138.10 GI:30911144
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 269155)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus chromosome 19, clone RP23-41B18
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 269155)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 269155)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 269155)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 5 (bases 1 to 269155)

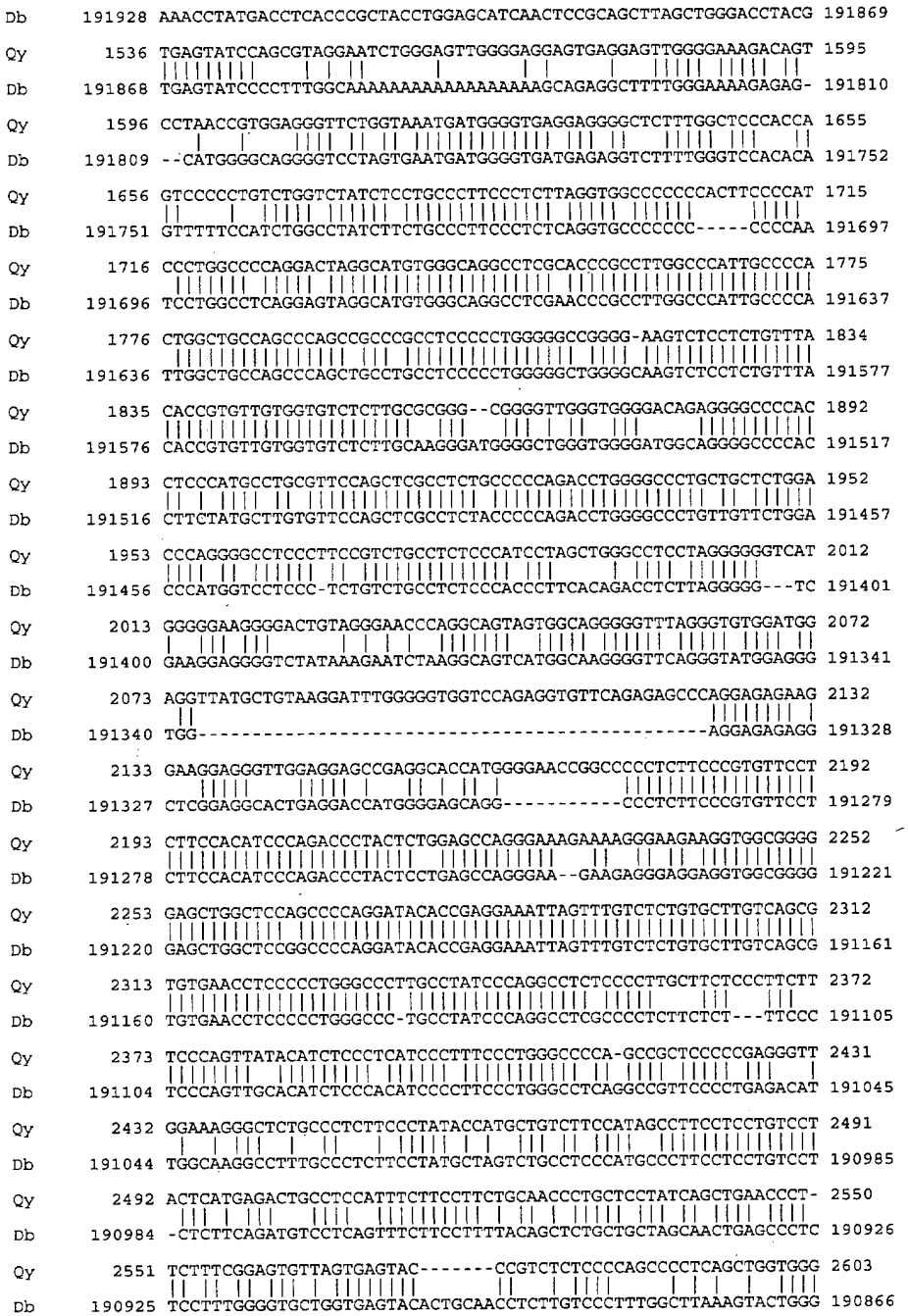
```
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
          Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
          Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
          Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
          Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
          Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
          Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
          Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
          Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., MacLean,C.,
          Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
          Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
          Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
          Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
          Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
          Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
          Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
          Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE     Direct Submission
JOURNAL   Submitted (01-OCT-2003) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On May 20, 2003 this sequence version replaced gi:30141982.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
          -------------- Project Information
              Center project name: L14003
              Center clone name: 41_B_18
          --------------.
FEATURES             Location/Qualifiers
     source          1. .269155
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="19"
                     /map="19"
                     /clone="RP23-41B18"
                     /clone_lib="RPCI-23 Female Mouse BAC"
     misc_feature    1. .6
                     /note="clone_boundary
                     clone_end:SP6
                     site:EcoRI"
     repeat_region   230. .253
                     /rpt_family="(TCTG)n"
     repeat_region   1705. .1731
                     /rpt_family="AT_rich"
     repeat_region   complement(1761. .1885)
                     /rpt_family="RSINE1"
     repeat_region   complement(2074. .2158)
                     /rpt_family="B1F"
     repeat_region   complement(2083. .2181)
                     /rpt_family="ID_B1"
     repeat_region   complement(2212. .2283)
                     /rpt_family="ID4"
     repeat_region   complement(2294. .2509)
                     /rpt_family="L1MB3"
     repeat_region   2510. .2536
                     /rpt_family="(TTTG)n"
     repeat_region   complement(2537. .2680)
                     /rpt_family="B1_MM"
     repeat_region   complement(2681. .2897)
                     /rpt_family="L1MB3"
     repeat_region   complement(2942. .3100)
                     /rpt_family="L1MB3"
     repeat_region   3168. .3403
                     /rpt_family="YREP_MM"
     repeat_region   complement(3237. .3444)
                     /rpt_family="ID_B1"
     repeat_region   3480. .3870
                     /rpt_family="MTD"
     repeat_region   complement(3884. .4011)
                     /rpt_family="RSINE1"
     repeat_region   4129. .4206
                     /rpt_family="L1MB3"
     repeat_region   complement(4427. .4480)
                     /rpt_family="MIR"
     repeat_region   4529. .4605
                     /rpt_family="PB1D7"
     repeat_region   4645. .4761
                     /rpt_family="ID_B1"
     repeat_region   4762. .4783
                     /rpt_family="B1F"
     repeat_region   10536. .10694
                     /rpt_family="B4"
     repeat_region   10701. .10843
                     /rpt_family="RSINE1"
     repeat_region   10869. .11015
                     /rpt_family="B1_MM"
     repeat_region   11044. .11144
                     /rpt_family="PB1D10"
     repeat_region   11489. .11563
                     /rpt_family="CT-rich"
     repeat_region   14157. .14251
                     /rpt_family="A-rich"
     repeat_region   15523. .15653
                     /rpt_family="URR1B"
     repeat_region   15714. .15791
                     /rpt_family="ID4"
     repeat_region   complement(17274. .17411)
                     /rpt_family="B1_MM"
     repeat_region   18732. .18881
                     /rpt_family="MT2B"
     repeat_region   18770. .18904
                     /rpt_family="B1_MM"
     repeat_region   complement(18942. .19087)

Query Match             33.6%;  Score 1709.4;  DB 10;  Length 269155;
Best Local Similarity   74.5%;  Pred. No. 0;
Matches 2733;  Conservative    0;  Mismatches  742;  Indels  193;  Gaps   39;

Qy       1207 CCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCCTCACCC 1266
              |||| ||| |||||||||||||||| ||||||     |||||||||||| || | |||||||||
Db     228433 CCAGAAGTCTTTCCCCTCCTCCACCCCCTCCATTTCCCATCTGGTCTCTCTCCCTCACCC 228492

Qy       1267 CCC-ACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCCGTGTG 1325
              ||| || ||||||||||||||||||||||||||||||||||||||||| |||||  | ||
Db     228493 CCCTTCCTAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGTTTCCCTATTTG 228552

Qy       1326 GGCCCGGG-CCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGGGATG 1384
              |||| ||| || | ||||| ||||| | |||||||   ||||||||||||||||||||||
Db     228553 GGCCTGGGCCCACCCTCATACTTCTTGCCCATCCTATTCACAGGGGACTCGTGGGGGATG 228612

Qy       1385 TTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCGCACA 1444
              ||||| ||||| |||||||||| ||||||||||||||||||||||||||||| |||||||||
Db     228613 TTAGCTTGCCTATGCACGGTGCTGTGGCACCTCCCTGCAGTGCCAGCTCTTAATCGCACA 228672

Qy       1445 GGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAG 1504
              || || ||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     228673 GGAGATCCAGGCCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCTGGAG 228732

Qy       1505 CACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGGAGTT 1564
              || |||||||||||||| |||||||||| ||||||||||     | | ||
Db     228733 CATCAACTCCGCAGCTTAGCTGGGACCTACGTGAGTATCCCCTTTGGCAA---------- 228782

Qy       1565 GGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAATGAT 1624
                ||||||  ||  | |  ||  |     |   |    |||| || || ||||||
Db     228783 --AGAGGAGGCAGAGGCTTTGGGGAAGAGAGAGCATGGGGCAGGGGTCCTAGTGAATGAT 228840

Qy       1625 GGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTGCCCT 1684
```

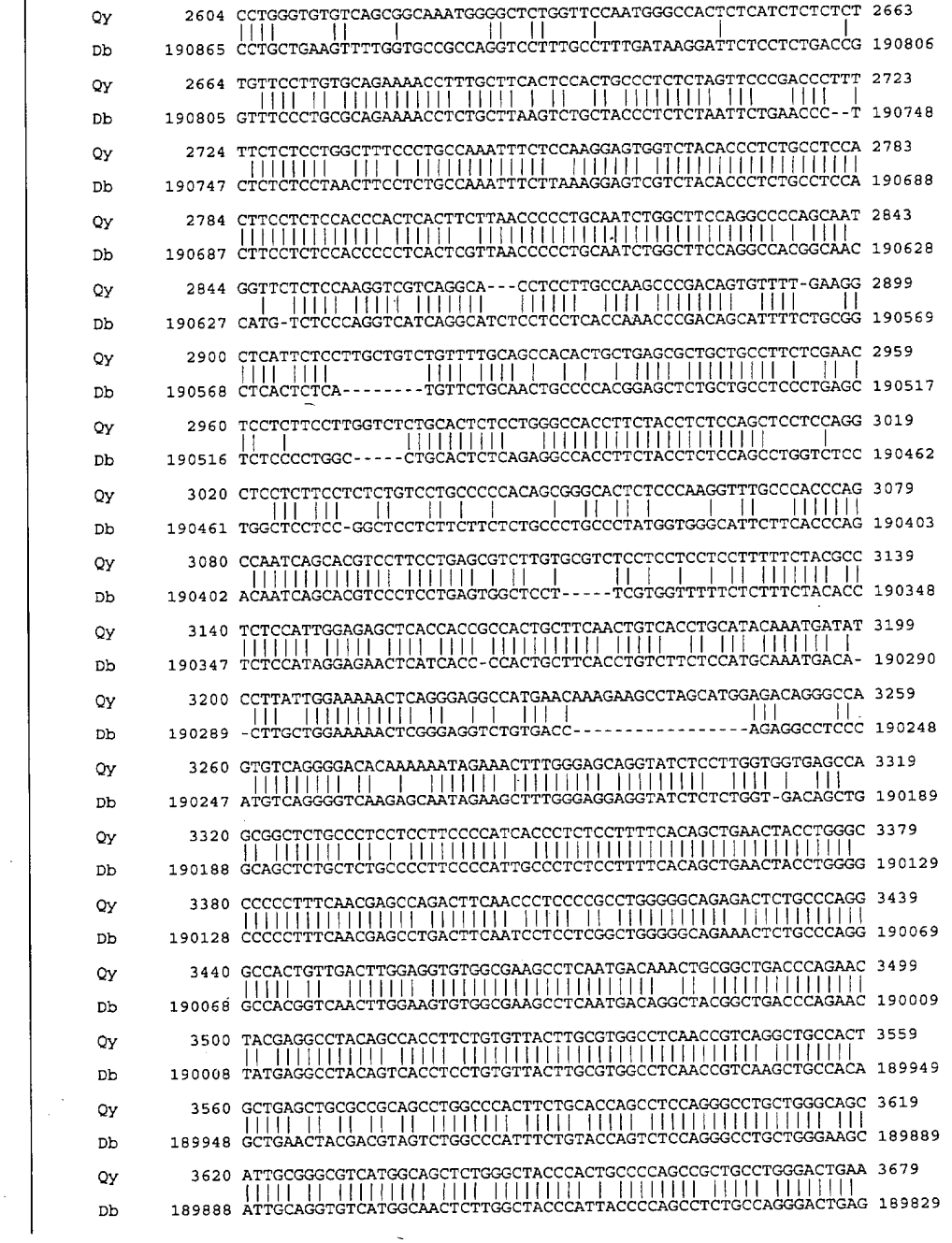

Db 228841 GGGGTGATGAGAGGCCCTTTGGGTCCACATAGCTTCTCCATCTGGCCTATCTTCTGCCCT 228900
Qy 1685 TCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTGGGCA 1744
Db 228901 TCCCTCTCAGGTGCCCCCC-------CCCAATCCTGGCCCCAGGAGTAGGCATGTGGGCA 228953
Qy 1745 GGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCCTCCC 1804
Db 228954 GGCCTCGAACCCGCCTTGGCCCATTGCCCCATTGGCTGCCAGCCCAGCTGCCTGCCTCCC 229013
Qy 1805 CCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCGGGCG 1864
Db 229014 CCTGGGGGCTGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCAAGGGTG 229073
Qy 1865 GGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTCTGCC 1924
Db 229074 GGGCTGGGTGGGGATGGAGGGGCCCCACCTTCTATGCTTGTGTTCCAGCTCGC--CTACC 229131
Qy 1925 CCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCTCCCA 1984
Db 229132 CCCAGACCTGGGGCCCTGTTGTTCTGGACCCATGGGCCTCCC-TCTGTCTGCCTCTCCCA 229190
Qy 1985 TCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGGCAGT 2044
Db 229191 TCCTTCACAGACCTCTTAGGGGG---TTGGAGGAGGGATCTACAAAGAATCTAGGGCAGT 229247
Qy 2045 AGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGGTCCA 2104
Db 229248 CATGGCAAGGGGTTCAGGGTATGGAGGGTGGTTGTGCCATAAGAATTTGAGGGTAGTTTA 229307
Qy 2105 GAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACCATGG 2164
Db 229308 GAGATGTTCACAGAGCCCAGGAGAGACGGAGAGAGGATCGGAGGCACTGAGG-ACCATGG 229366
Qy 2165 GGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAGCCAG 2224
Db 229367 GGAACAGG-CCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCCTGAGCCAG 229425
Qy 2225 GGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGA 2284
Db 229426 GG-AAGAAAAAGGGAGGAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCGAGGA 229484
Qy 2285 AATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTATCCC 2344
Db 229485 AATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCAGTGCCTATCCC 229544
Qy 2345 AGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTTTCCC 2404
Db 229545 AGGCCTCGCCCCTCTTCTCT---TTCCCTCCCAGTTGCACATCTCCCACATCCCCTTCCC 229601
Qy 2405 TGGGCCCCA-GCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACCATG 2463
Db 229602 TGGGCCTCAGGCCGTTCCCCTGAGACATTGGCAAGGACTTTGCCCTCTTCCGATGCTCCT 229661
Qy 2464 CTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTTCTG 2523
Db 229662 CTGCCTCCCATGCCCTCCCTCCTGTCCT-CTCTTTAGATTTCCTCAGTTTCTTCCTT-TA 229719
Qy 2524 CAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTCTCC 2583
Db 229720 CAGCTCTGCTTCTAGGAACTGAGTTCTCTTCTTTGGG-----GTAAGTACATTGCTATCT 229774
Qy 2584 CCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCAATG 2643
Db 229775 CTTGTCCCTTTGGCCTGTTGAGGAGGAGTTTTGGTGTCAGCAGGTCCCTTTGCCTTTGAT 229834
Qy 2644 GGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTGCCC 2703
Db 229835 ATGGACTCTCCTCTGGCTGGTTTTCCT-GCACAGAAAACCTCTACTTGAGTCTGCTACCC 229893
Qy 2704 TCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGAGTG 2763
Db 229894 TCTCTAACTC----GACCCTTTCTCTCCTAACTTCCTCTGCCAAATTTCTTGAAGGAGTC 229949
Qy 2764 GTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAATCT 2823
Db 229950 GTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTGTTAACCCCCTGCAATCT 230009
Qy 2824 GGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAGCCC 2883
Db 230010 GGCTTCCAGGCCACGGCAAT-CATCTCTCCCAGGTCATCAGGCATCTCCTTCTCA----C 230064
Qy 2884 GACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAGCGC 2943
Db 230065 CAAACCCGACAGCATTTTCTGCGGCTCACTCTCATGTTCTGCAACTGCCCTGCGGAGCTC 230124
Qy 2944 TGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTACCT 3003
Db 230125 TGCTGCCGTCCTGAGCTCTTC-----CCTGGCCTGCACTCTCAGAGGCCACCTTCTACCT 230179
Qy 3004 CTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTCCCA 3063
Db 230180 CTCCAGCCTGGTCTCCTGGCTCCTCCG-GCTCCTCTTCTTCTCTGCCCTGCCCTGTGGTG 230238
Qy 3064 AGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTCCTC 3123
Db 230239 GGCATTCTCCACCCAGACAGTCAGCACGTCCCACCCGAGTGGCTCCTTCGTGGTTTTCTC 230298
Qy 3124 CTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTCACC 3183
Db 230299 ------TCTCTGCACTTCTCCATAGGAGAACTCATCACC-CCACTGCTTCACCTGTCTCC 230351
Qy 3184 TGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCCTAG 3243
Db 230352 TCCATGCAAATGATACACTTGCTGGAAAACCT---------------------------- 230383
Qy 3244 CATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTATCT 3303
Db 230384 -------------CCCATGTCAGGGGTCAAGAGCGATAGAAGCTTTGGGAGGAGGTATCT 230430
Qy 3304 CCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACA 3363
Db 230431 CTCTGCTGGCAAGCTGGCAGCTCTGCTCTGCTCCTTCCCCATTGCCCTCTCCTTTTCACA 230490
Qy 3364 GCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGC 3423
Db 230491 GCTGAACTACCTGGGGCCCCCTTTCAACGAGCCTGACTTCAATCCTCCTCGACTGGGGGC 230550
Qy 3424 AGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACT 3483
Db 230551 AGAAACTCTGCCCAGGGCCACGGTCAACTTGGAAGTGTGGCGAAGCCTCAATGACAGGCT 230610
Qy 3484 GCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAA 3543
Db 230611 GCGGCTGACCCAGAACTATGAGGCGTACAGTCACCTCCTGTGTTACTTGCGTGGCCTCAA 230670
Qy 3544 CCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCA 3603
Db 230671 CCGTCAGGCTGCCACAGCTGAACTCCGACGTAGCCTGGCCCACTTCTGTACCAGCCTCCA 230730
Qy 3604 GGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCC 3663
Db 230731 GGGCCTGCTGGGCAGCATTGCAGGTGTCATGGCGACGCTTGGCTACCCACTGCCCCAGCC 230790
Qy 3664 GCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAA 3723
Db 230791 TCTGCCAGGGACTGAGCCAGCCTGGGCCCCTGGCCCTGCCCACAGTGACTTCCTCCAGAA 230850
Qy 3724 GATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGA 3783
Db 230851 GATGGATGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTATGGCGTTCAGCCAAGGA 230910
Qy 3784 CTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGG 3843
Db 230911 CTTCAACCGGCTTAAGAAGAAGATGCAGCCTCCAGCAGCTTCAGTCACCCTGCACTTGGA 230970



RESULT 10
AC135823/c
LOCUS AC135823 258710 bp DNA linear HTG 20-NOV-2002
DEFINITION Rattus norvegicus clone CH230-42H20, *** SEQUENCING IN PROGRESS ***, 6 unordered pieces.
ACCESSION AC135823
VERSION AC135823.2 GI:25138863
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 258710)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 258710)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 258710)

```
AUTHORS     Rat Genome Sequencing Consortium.
TITLE       Direct Submission
JOURNAL     Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:24270628.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: KCLS
                Center clone name: CH230-42H20
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 212150 bases at least Q40
                Consensus quality: 217133 bases at least Q30
                Consensus quality: 220786 bases at least Q20
                Estimated insert size: 230118; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    98082: contig of 98082 bp in length
            *    98083    98182: gap of unknown length
            *    98183   119083: contig of 20901 bp in length
            *   119084   119183: gap of unknown length
            *   119184   210852: contig of 91669 bp in length
            *   210853   210952: gap of unknown length
            *   210953   255563: contig of 44611 bp in length
            *   255564   255663: gap of unknown length
            *   255664   257428: contig of 1765 bp in length
            *   257429   257528: gap of unknown length
            *   257529   258710: contig of 1182 bp in length.
FEATURES             Location/Qualifiers
     source          1..258710
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-42H20"
     misc_feature    4922..5734
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:BH348759"
     misc_feature    16913..17451
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BH348731"
     misc_feature    23010..24091
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    98183..99248
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    116266..119083
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    210953..212523
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    212785..214316
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

  Query Match             32.1%;  Score 1632.6;  DB 2;  Length 258710;
  Best Local Similarity   71.1%;  Pred. No. 0;
  Matches 2996;  Conservative    0;  Mismatches  959;  Indels  258;  Gaps   53;

Qy        704 GAGCACTAAATCTCCTTCAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACAT 763
Db     192673 GAATGCTATAGCTCCTTCAGT-TAACCCTATCCTGCCCTCACAACACCCAGGAACTATAT 192615

Qy        764 CCTAGAATATACCCAAAACTAAGCCCCATAAGGCAGCCCGACCCTAG-----TGGTCTAA 818
Db     192614 CCTAGTATATACCCTCCACTTGTCACTCCACAAACCGGTTATCTTGGACCCTTGGCCTGA 192555

Qy        819 CCCTATACCTTGCTTCCTATGGGTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCC 878
Db     192554 CCTCAGCTGTGGAACACTACTCCTATCTCCCAGAGCCGCTGCCTCTCCTCTCACTCTTCC 192495

Qy        879 TCCCTTAGAGCTGACTGTGCTCAGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCT 938
Db     192494 AC----AGGGTTCACTGCTCACAGCTTGCCAGAGCTAACATATGTCATATTCCCACCCTT 192439

Qy        939 GACTCCCCTCAAGCTGCAGTGGGACTGGAAGACTGGCAGGA-AGCTAGGGTACAACTGGA 997
Db     192438 TAGTTCCT----GCTGCAATGGAACTGGAAGACTTGGAGAAGAGGCTGGACACAACTGGA 192383

Qy        998 ACACAGGCAGGTCGAC-CTGCAGTCCCTAGGCCTGGCCCCGTCCCTCCATGTACACACAT 1056
Db     192382 ACGTCAGCATATTGACACAAAAGTCTCTAAAGCCAGGCCTGGCCCTCTGAGTAGACACGT 192323

Qy       1057 ATACATGTTGGCACACACACAGTGGCACACATGCCAAAGACTCTCTCAGCTGACACACAG 1116
Db     192322 GCACACATATTTGGCACATATAGGGAAGACATGCCAAAGACTCGGCTGATAG-------- 192271

Qy       1117 ATCCATTCTCAAGTATCTACTGATAGACACTCATGCGTGCCAAGTCCTCATCCTCAAACA 1176
Db     192270 -----CAGTCGAGAATCTACTGACAGACACACATGCCAGGACTGTTCTCGCAGACAGACC 192216

Qy       1177 TACACATGCCTCTCTTTCTCTCCCGTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTC 1236
Db     192215 TCCACATACA------------CTCTTGACCACAAGTCTTTCCCCTCCTCCATCCCCTC 192169

Qy       1237 TGCCTCCCATCTGGTGTCCCACCCTCACCCCCCACCCAGCCCAAGGTGGGGACAGACACC 1296
Db     192168 CACCTTCCATCTGGTCTCTCTCCCTCACCCCCTTCCTAGCCCAAGGTGGGGACAGACACC 192109

Qy       1297 TGAGGGGCTGCCAGCTGCTTCCCCGTGTGGGCCCGGG-CCGCGCTCATGCTTCTCGTCCA 1355
Db     192108 TGAGGGGCTGCCAGCTGTTTCCCCACGTGGGCCTGGGCCCACCCTCATACTTCTTGCCCA 192049

Qy       1356 TCCTGCCCACAGGGGACTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACC 1415
Db     192048 TCCTATCCACAGGGGACTCGTGGGGGATGTTAGCTTGCCTCTGCACGGTGCTGTGGCACC 191989

Qy       1416 TCCCTGCAGTGCCAGCTCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGA 1475
Db     191988 TCCCTGCAGTGCCAGCTCTTAATCGAACAGGAGATCCAGGCCCTGGCCCCTCCATCCAGA 191929

Qy       1476 AAACCTATGACCTCACCCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATG 1535
```





```
Qy       3680 CCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGG 3739
Db     189828 CCAGCCTGGGCCCCTGGCCCTGCCCACAGCGACTTCCTCCAGAAGATGGATGACTTCTGG 189769

Qy       3740 CTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAG 3799
Db     189768 CTGCTGAAGGAGCTGCAGACCTGGCTATGGCGTTCAGCCAAAGACTTCAACCGGCTTAAG 189709

Qy       3800 AAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGA 3859
Db     189708 AAGAAGATGCAACCTCCAGCAGCTTCAGTCACCCTGCACTTGGAGGCCCATGGTTTCTGA 189649

Qy       3860 CTTCTGACCTT-----CTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTTTGTGAGA 3914
Db     189648 TCTCTGACCCTTCACCCCCACACCTTCAGCCCCAGTCAACTGTGCTTCCATTGTGGGGTG 189589

Qy       3915 GCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGGCCC 3974
Db     189588 AAAAATCCTCTATACCAATATTTGTTGAGCCAGGAGCCAGACCATAGTAG---CTCTACC 189532

Qy       3975 TTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCAAAG 4034
Db     189531 CTCCTCGCCTTGGATGGGGGTGTAACTCAATAAGCCCTCTAGCTCCTTGA---------- 189482

Qy       4035 GTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGAAGT 4094
Db     189481 GTCCTAGGGTCTGGGGAAAAGGTACAGTGGGCAC----------------CAAGAGGT 189440

Qy       4095 CATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCTGCT 4154
Db     189439 AATGCCTGAGGGAGTCGAA---TGGCTCAGTGGATAGAAAGGCCCTCTCTGCCTCCTGCT 189383

Qy       4155 TCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACAT---------- 4203
Db     189382 TTCTGCTTGCCACTT-GCCAGTAACCACTCAGTTCCTCATGTGGCACTTGCAGAAACATA 189324

Qy       4204 -CTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
Db     189323 CCTGGAGGGCATGGGT--GGGGGCCACTGTAGCATGTGCCTTTCTGGGATAAGGCCCCTT 189266

Qy       4263 GGCTGCCC-CACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACAT 4321
Db     189265 GGCTTCCCTCCCTCTCCTTGGATGGGTGTTGCTTCCCTACCCCAATA----ACTACATAT 189210

Qy       4322 CCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGC 4381
Db     189209 CCAGTTCAGGAAACAAA---------GTGACCAGTCTAAAGAAGAGTGTGATAGAAGTGG 189159

Qy       4382 AGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACA 4441
Db     189158 AGGGGTAGGGTCTGTGCTGGAGGTGGCCTAGAAACCAGAAGATGTAATTCTAAAAG-AGA 189100

Qy       4442 GGGGCAGGGACAGACCAGAC----CAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAAC 4497
Db     189099 AGGGCAGGGACAGAACAGACCCAGCAGTCACCAAGACAGTGGTGGCACAGGGCGGAAAAC 189040

Qy       4498 AAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTT-CCAGTATTACGGTGCCT 4556
Db     189039 AAAAGCCATGTTGAGCATTAAGACCTTGCCTTGAATTTTCTTGCCAGCATCTTGGTACCT 188980

Qy       4557 CTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAG 4616
Db     188979 CCTCTCTGCCCCCTTTCCCAAGGTATCTGTGGTTGCCTGGGCTGGGGAGGGCAGCCATAG 188920

Qy       4617 CCACA-CCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGG 4675
Db     188919 CCATAGCCACAGGGTTTCCTGAAAGTTTACATTGCAGTAGCATTGTGGGGT---AAGGGG 188863

Qy       4676 CAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTC-TAAGTGTGTTGTAT 4734
Db     188862 GATGGCTCCCCAGGCTCTATCCCCAGCCCCACCCACTCATGACTCTTAAGTTTGTTGTAT 188803

Qy       4735 TAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATT 4794
Db     188802 TAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATT 188743

Qy       4795 CTTGTATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCC 4854
Db     188742 CTTGTATTAACAAATAAAATGCTTGCCCCAGAACTTGGACTCTGTCTAGCTTGGTAATCT 188683

Qy       4855 TCCTGGTGCTCAT 4867
Db     188682 TCCTGGGGGCCAT 188670
```

RESULT 11
BD132824
LOCUS BD132824 1710 bp DNA linear PAT 18-SEP-2002
DEFINITION Cardiotrophin-like cytokine.
ACCESSION BD132824
VERSION BD132824.1 GI:23227769
KEYWORDS JP 2002507125-A/1.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 1710)
AUTHORS Shi,Y. and Ruben,S.M.
TITLE Cardiotrophin-like cytokine
JOURNAL Patent: JP 2002507125-A 1 05-MAR-2002;
HUMAN GENOME SCIENCES INC
COMMENT PN JP 2002507125-A/1
PD 05-MAR-2002
PF 29-JUN-1998 JP 1999505681
PR 30-JUN-1997 US 60/051311
PI YANGGU SHI,STEVEN M RUBEN
PC C07K14/435,C07K14/52,C07K16/24,C12N1/21,C12N15/19,C12N15/64 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 46. .720
FT sig peptide 46. .126
FT mat peptide 127. .720.
FEATURES Location/Qualifiers
source 1. .1710
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN

Query Match 28.8%; Score 1463.4; DB 6; Length 1710;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
Db        227 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 286

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
Db        287 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 346

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
Db        347 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 406

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
Db        407 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 466

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
Db        467 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 526

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
Db        527 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 586
```

```
Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 646

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 706

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 766

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 826

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 886

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 946

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1006

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1066

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1126

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1186

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1246

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1306

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1366

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1426

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1486

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1546

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1606

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1666

Qy       4803 AACAAATAAAATGCTTGCCCCAGAA 4827
              |||||||||||||||||||||||||
Db       1667 AACAAATAAAATGCTTGCCCCAGAA 1691
```

RESULT 12
BC012939

LOCUS BC012939 1736 bp mRNA linear PRI 04-OCT-2003
DEFINITION Homo sapiens cardiotrophin-like cytokine, mRNA (cDNA clone MGC:21195 IMAGE:4453813), complete cds.
ACCESSION BC012939
VERSION BC012939.1 GI:15277894
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1736)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1736)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 28 Row: b Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7019350.

FEATURES Location/Qualifiers
source 1. .1736
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```
                     /clone="MGC:21195 IMAGE:4453813"
                     /tissue_type="Kidney, hypernephroma"
                     /clone_lib="NIH_MGC_89"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1736
                     /gene="CLC"
                     /note="synonym: BSF3"
                     /db_xref="LocusID:23529"
                     /db_xref="MIM:607672"
     CDS             73. .750
                     /codon_start=1
                     /product="cardiotrophin-like cytokine"
                     /protein_id="AAH12939.1"
                     /db_xref="GI:15277895"
                     /db_xref="LocusID:23529"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
     misc_feature    313. .693
                     /note="CNTF; Region: Ciliary neurotrophic factor"
                     /db_xref="CDD:pfam01110"
ORIGIN

  Query Match             28.8%;  Score 1463.4;  DB 9;  Length 1736;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1464;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 313

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 373

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 433

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 493

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 553

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        554 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 613

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        614 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 673

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        674 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 733

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        734 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 793

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        794 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 853

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        854 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 913

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        914 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 973

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        974 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1033

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1093

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1153

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1213

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1273

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1333

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1393

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1453

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1513

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1573

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1574 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1633

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1634 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1693

Qy       4803 AACAAATAAAATGCTTGCCCCAGAA 4827
             ||||||||||||||||||||||||||
Db       1694 AACAAATAAAATGCTTGCCCCAGAA 1718


RESULT 13
AF172854
LOCUS       AF172854               1689 bp    mRNA    linear   PRI 10-SEP-1999
DEFINITION  Homo sapiens cardiotrophin-like cytokine CLC mRNA, complete cds.
ACCESSION   AF172854
VERSION     AF172854.1  GI:5852980
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1689)
  AUTHORS   Shi,Y., Wang,W., Yourey,P.A., Gohari,S., Zukauskas,D., Zhang,J.,
            Ruben,S. and Alderson,R.F.
  TITLE     Computational EST database analysis identifies a novel member of
```

```
          the neuropoietic cytokine family
  JOURNAL   Biochem. Biophys. Res. Commun. 262 (1), 132-138 (1999)
  MEDLINE   99382254
   PUBMED   10448081
REFERENCE   2  (bases 1 to 1689)
  AUTHORS   Shi,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-1999) Molecular Biology, Human Genome Science,
            Inc., 9410 Key West Avenue, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .1689
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q13.3"
     CDS             46. .723
                     /codon_start=1
                     /product="cardiotrophin-like cytokine CLC"
                     /protein_id="AAD54284.1"
                     /db_xref="GI:5852981"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match           28.7%;  Score 1461.4;  DB 9;  Length 1689;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1462;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 286

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 346

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 406

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 466

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 526

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 586

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 646

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 706

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 766

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 826

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 886

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 946

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1006

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1066

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1126

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1186

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1246

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1306

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1366

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1426

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1486

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1546

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1606

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1666

Qy       4803 AACAAATAAAATGCTTGCCCCAG 4825
             |||||||||||||||||||||||
Db       1667 AACAAATAAAATGCTTGCCCCAG 1689


RESULT 14
AX205024
LOCUS       AX205024                881 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 1 from Patent WO0155172.
ACCESSION   AX205024
VERSION     AX205024.1  GI:15394259
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elson,G., Gauchat,J.F., Plun-Favreau,H., Chevalier,S. and Gascan,H.
```

```
  TITLE     Isolated complex comprising a nnt-1 protein and in addition at
            least a clf-1 protein and/or a scntfr_g(a) protein
  JOURNAL   Patent: WO 0155172-A 1 02-AUG-2001;
            PIERRE FABRE MEDICAMENT (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE
            LA RECHERCHE MEDICALE (INSERM) (FR)
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             174. .851
                     /note="unnamed protein product; NNT-1"
                     /codon_start=1
                     /protein_id="CAC60175.1"
                     /db_xref="GI:15394260"
                     /db_xref="REMTREMBL:CAC60175"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match           10.3%;  Score 525.4;  DB 6;  Length 881;
  Best Local Similarity  99.8%;  Pred. No. 9e-114;
  Matches  526;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 414

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 474

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 534

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 594

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 654

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 714

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 774

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 834

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 3889
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 881


RESULT 15
AX205042
LOCUS       AX205042                881 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 1 from Patent WO0155219.
ACCESSION   AX205042
VERSION     AX205042.1  GI:15394277
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Elson,G. and Gauchat,J.F.
  TITLE     Scscntfr/nnt-1 fusion protein
  JOURNAL   Patent: WO 0155219-A 1 02-AUG-2001;
            PIERRE FABRE MEDICAMENT (FR)
FEATURES             Location/Qualifiers
     source          1. .881
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             174. .851
                     /note="unnamed protein product; NNT-1"
                     /codon_start=1
                     /protein_id="CAC60178.1"
                     /db_xref="GI:15394278"
                     /db_xref="REMTREMBL:CAC60178"
                     /translation="MDLRAGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTY
                     DLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLR
                     LTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQ
                     PLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTL
                     HLGAHGF"
ORIGIN

  Query Match           10.3%;  Score 525.4;  DB 6;  Length 881;
  Best Local Similarity  99.8%;  Pred. No. 9e-114;
  Matches  526;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 414

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 474

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 534

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 594

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 654

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 714

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 774

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 834

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 3889
             |||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 881


Search completed: August 15, 2004, 23:47:32
Job time : 19051.8 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 15:26:00 ; Search time 11579.5 Seconds
(without alignments)
13118.770 Million cell updates/sec

Title: US-09-931-704-3
Perfect score: 5087
Sequence: 1 aacctgcgagtgggcctggc.....cctttgctaagtcttcctca 5087

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	923.2	18.1	1025	9	AL570325	AL570325 AL570325
	2	860.2	16.9	1028	9	AL543945	AL543945 AL543945
c	3	755.4	14.8	773	14	CA430870	CA430870 UI-H-FL1-
c	4	663.4	13.0	681	13	BU631799	BU631799 UI-H-FL0-
c	5	656.6	12.9	681	10	AW978561	AW978561 EST390670
c	6	637.4	12.5	679	14	CA306071	CA306071 UI-H-FT1-
	7	634.4	12.5	744	13	BX092291	BX092291 BX092291
c	8	618.2	12.2	675	14	CA748807	CA748807 UI-H-FT1-
	9	612	12.0	691	10	BF213570	BF213570 601845370
	10	611.6	12.0	831	28	CC138171	CC138171 NDL.3C1.T
c	11	588	11.6	793	14	CK032835	CK032835 AGENCOURT
c	12	578	11.4	606	14	CA309862	CA309862 UI-H-FT1-
c	13	577.8	11.4	763	14	CF271924	CF271924 AGENCOURT
c	14	576.8	11.3	801	14	CF271930	CF271930 AGENCOURT
c	15	572.8	11.3	800	14	CF271926	CF271926 AGENCOURT
	16	567.4	11.2	915	10	BF035982	BF035982 601457860
	17	540.2	10.6	806	14	CK130211	CK130211 AGENCOURT
	18	537.2	10.6	787	14	CK130210	CK130210 AGENCOURT
	19	527	10.4	788	14	CF271927	CF271927 AGENCOURT
	20	524	10.3	912	13	BX389473	BX389473 BX389473
	21	512.4	10.1	514	10	BF058969	BF058969 7k36e10.x
	22	511	10.0	752	14	CF271925	CF271925 AGENCOURT
	23	511	10.0	804	14	CF271929	CF271929 AGENCOURT
	24	502.2	9.9	810	14	CK130212	CK130212 AGENCOURT
	25	495.4	9.7	662	29	AY409170	AY409170 Homo sapi
	26	490.6	9.6	662	29	AY409171	AY409171 Pan trogl
	27	488.6	9.6	1157	13	BQ940483	BQ940483 AGENCOURT
c	28	487	9.6	487	10	BE463575	BE463575 hw25d05.x
c	29	479.4	9.4	493	9	AI040033	AI040033 ox97h03.x
c	30	478	9.4	495	14	CD369579	CD369579 UI-H-FT1-
	31	471.4	9.3	853	12	BG437538	BG437538 602489277
	32	470.4	9.2	887	13	BQ948158	BQ948158 AGENCOURT
c	33	465.4	9.1	467	10	BF002167	BF002167 7h23c10.x
c	34	458.4	9.0	510	14	CB045938	CB045938 NISC_gf01
c	35	452.4	8.9	454	10	BF513712	BF513712 UI-H-BW1-
c	36	450.4	8.9	457	9	AW044541	AW044541 wy73d06.x
c	37	446.2	8.8	462	9	AI916548	AI916548 wa26h07.x
	38	437.2	8.6	465	9	AA430602	AA430602 zw21a11.r
c	39	429.4	8.4	432	9	AA430517	AA430517 zw21a11.s
	40	422.4	8.3	432	9	AA278936	AA278936 zs78d06.r
c	41	422.2	8.3	433	9	AI765382	AI765382 wi74d07.x
	42	421.8	8.3	662	29	AY409172	AY409172 Mus muscu
c	43	417.6	8.2	766	14	CF271928	CF271928 AGENCOURT
c	44	407.4	8.0	409	9	AA283037	AA283037 zt16d01.s
c	45	405	8.0	411	9	AI752562	AI752562 cn17d05.y

ALIGNMENTS

RESULT 1
AL570325/c
LOCUS AL570325 1025 bp mRNA linear EST 31-MAY-2003
DEFINITION AL570325 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI004YM15 3-PRIME, mRNA sequence.
ACCESSION AL570325
VERSION AL570325.2 GI:31291747
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1025)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12926520.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6127.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI004AG08NP1&cluster=6127.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

```
          http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0DI004AG08NP1.
FEATURES             Location/Qualifiers
     source          1. .1025
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI004YM15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            18.1%;  Score 923.2;  DB 9;  Length 1025;
  Best Local Similarity  95.8%;  Pred. No. 1.6e-189;
  Matches  965;  Conservative   18;  Mismatches   19;  Indels    5;  Gaps    4;

Qy       3812 CCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTC 3871
             :|:|| :::||||||:||||||||||||||||||||||||||||||||  |||||||:
Db       1006 SCYCCCRSMGCTGCAKTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTMTTGACCTTY 947

Qy       3872 TCCTCTTCGCTC-CCCCTTCAAACCCTGCTCCCACTTTGTGAGAGCCAGCCCTGTATG-C 3929
             || |:|:||||| |||:|||||||||||||||||||||||||||||||||||||||||| |
Db        946 TCTTYTKCGCTCYCCCYTTCAAACCCTGCTCCCACTTTGTGAGAGCCAGCCCTGTATGCC 887

Qy       3930 CAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGGCCCTTTCCTGGACCGGCT 3989
             | ||||||||||||||||||||||||||||||||:|||||||:|||||||||||||||||
Db        886 CTACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCYTCTGGCCYTTTCCTGGACCGGCT 827

Qy       3990 GGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCAAAGGTCTACCGAGCTGGG 4049
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        826 GGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACCTCCCAAAGGTCTACCGAGCTGGG 767

Qy       4050 GAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGAAGTCATGCTCGAGGGAGT 4109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 GAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTACAGGAAGTCATGCTCGAGGGAGT 707

Qy       4110 GTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCTGCTTCTTGCCTACCACTT 4169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        706 GTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGCCTCCTGCTTCTTGCCTACCACTT 647

Qy       4170 GGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGGCAGGGGTTGAGGGGCCAC 4229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 GGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGGCAGGGGTTGAGGGGCCAC 587

Qy       4230 CACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTG 4289
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 CACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTG 527

Qy       4290 TTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAA 4349
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 TTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAA 467

Qy       4350 TTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCC 4409
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TTCTACAC-AAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCC 408

Qy       4410 TATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCA 4469
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 TATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCA 348

Qy       4470 GGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTC 4529
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 GGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTC 288

Qy       4530 GAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGG 4589
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 GAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGG 228

Qy       4590 TTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATG 4649
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 TTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATG 168

Qy       4650 CAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCC 4709
             ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db        167 CAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCSCCAGCCCCACCC 108

Qy       4710 ACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTATAGA 4769
             ||||||||||||||||:||||||||||||||||||||||||||||||||||||||| ||||
Db        107 ACTCATGACTCTAAGWGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTAATAGA 48

Qy       4770 TGATA--TTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAAT 4814
             ||    ||  | |||||||| |||| ||:|     |:|||:||||:
Db         47 TGNANATTTANTNGCAGAATTACTATACTWGNNMANAMAAAHAAAAW 1


RESULT 2
AL543945
LOCUS       AL543945               1028 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL543945 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI004YM15 5-PRIME, mRNA sequence.
ACCESSION   AL543945
VERSION     AL543945.2  GI:31265790
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1028)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12876424.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 6127.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI004AG08QP1&cluster=6127.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI004AG08QP1.
FEATURES             Location/Qualifiers
     source          1. .1028
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI004YM15"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            16.9%;  Score 860.2;  DB 9;  Length 1028;
  Best Local Similarity  99.2%;  Pred. No. 8.2e-176;
  Matches  896;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    3;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 185

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             |||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
Db        186 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAA--CTCAATGACAAAC 243

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        244 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 303

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 363

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 423

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 483

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 543

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 603

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 663

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 723

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 783

Qy       4023 CACCTCCCAAAGGTCTACCGAGCT-GGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGT 4081
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        784 CACCTCCCAAAGGTCTACCGAGCTGGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGT 843

Qy       4082 TTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTC 4141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        844 TTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTC 903

Qy       4142 ATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCAC 4201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 ATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCAC 963

Qy       4202 ATCTGGAGGGCA-GGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||| |||
Db        964 ATCTGGAGGGCAGGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCTT 1023

Qy       4261 TTG 4263
              |||
Db       1024 TGG 1026


RESULT 3
CA430870/c
LOCUS       CA430870                773 bp    mRNA    linear   EST 07-NOV-2002
DEFINITION  UI-H-FL1-bge-c-03-0-UI.s1 NCI_CGAP_FL1 Homo sapiens cDNA clone
            UI-H-FL1-bge-c-03-0-UI 3', mRNA sequence.
ACCESSION   CA430870
VERSION     CA430870.1  GI:24793596
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 773)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .773
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FL1-bge-c-03-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FL1"
                     /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                     (Pharmacia) with a modified polylinker; Site_1: EcoR I;
                     Site_2: Not I; NCI_CGAP_FL1 is a normalized cDNA library
                     derived from a pool of mRNA obtained from 4 cell lines
                     from grade III chondrosarcoma tissues. The library was
                     constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site.  Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     GAGGTCGGTG. The cell lines were provided by Dr. James
                     Martin from the University of Iowa.
                     TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
                     TAG_LIB=UI-H-FL1
                     TAG_SEQ=GAGGTCGGTG"
ORIGIN

  Query Match           14.8%;  Score 755.4;  DB 14;  Length 773;
  Best Local Similarity  99.9%;  Pred. No. 4.5e-153;
  Matches  756;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4071 TCCTGTCCTGTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTG 4130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 TCCTGTCCTGTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTG 714

Qy       4131 CAGAGGCGCTCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCC 4190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CAGAGGCGCTCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCC 654

Qy       4191 TCAGGTGGCACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGG 4250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCAGGTGGCACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGG 594

Qy       4251 GTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATC 4310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 GTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATC 534

Qy       4311 ACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAG 4370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 ACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAG 474

Qy       4371 ATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATA 4430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 ATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATA 414

Qy       4431 CTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGT 4490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 CTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGT 354

Qy       4491 GGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACG 4550
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        353 GGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTATG 294

Qy       4551 GTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAA 4610
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAA 234

Qy       4611 CCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAG 4670
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 CCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAG 174

Qy       4671 GGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTT 4730
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTT 114

Qy       4731 GTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTC 4790
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 GTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTC 54

Qy       4791 TATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAA 4827
                |||||||||||||||||||||||||||||||||||
Db          53 TATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAA 17
```

```
RESULT 4
BU631799/c
LOCUS       BU631799                681 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  UI-H-FL0-bdp-h-13-0-UI.s1 NCI_CGAP_FL0 Homo sapiens cDNA clone
            UI-H-FL0-bdp-h-13-0-UI 3', mRNA sequence.
ACCESSION   BU631799
VERSION     BU631799.1  GI:23299054
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 681)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FL0-bdp-h-13-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FL0"
                     /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                     (Pharmacia) with a modified polylinker; Site_1: EcoR I;
                     Site_2: Not I; NCI_CGAP_FL0 is a cDNA library derived from
                     a pool of mRNA obtained from 4 cell lines from grade III
                     chondrosarcoma tissues. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GAGGTCGGTG. The cell line
                     was provided by Dr James Martin from University of Iowa.
                     TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
                     TAG_LIB=UI-H-FL0
                     TAG_SEQ=GAGGTCGGTG"
ORIGIN
```

```
  Query Match           13.0%;  Score 663.4;  DB 13;  Length 681;
  Best Local Similarity  99.8%;  Pred. No. 4.4e-133;
  Matches  664;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4163 ACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGGCAGGGGTTGAG 4222
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 ACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGGCAGGGGTTGAG 622

Qy       4223 GGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGG 4282
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        621 GGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGG 562

Qy       4283 ATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATG 4342
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 ATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATG 502

Qy       4343 GTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGA 4402
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 GTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGA 442

Qy       4403 GGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACC 4462
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 GGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACC 382

Qy       4463 AGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACC 4522
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 AGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACC 322

Qy       4523 TTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTAT 4582
                |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 TTGCCTCGAATTGTCTTCCAGTATTATGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTAT 262

Qy       4583 CTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTT 4642
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 CTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTT 202

Qy       4643 TACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGC 4702
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGC 142

Qy       4703 CCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATT 4762
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 CCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATT 82

Qy       4763 TATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCC 4822
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 TATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCC 22

Qy       4823 CAGAA 4827
                |||||
Db          21 CAGAA 17
```

```
RESULT 5
AW978561/c
LOCUS       AW978561                681 bp    mRNA    linear   EST 02-JUN-2000
DEFINITION  EST390670 MAGE resequences, MAGP Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW978561
VERSION     AW978561.1  GI:8169828
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 681)
```

```
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
            Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 394
            Seq primer: Forward.
FEATURES             Location/Qualifiers
     source          1. .681
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGP"
                     /note="Vector: pBluescriptSKm"
ORIGIN

  Query Match           12.9%;  Score 656.6;  DB 10;  Length 681;
  Best Local Similarity  99.3%;  Pred. No. 1.3e-131;
  Matches  670;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;

Qy        190 TGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCA 249
              | |||||||||||||||| |||||||| |||||||||||||||||||||||||||||||
Db        681 TAATACCTAAACCGACCAATTCACAGCCTTCCAACTCACCCTCTGCCTGCCCAGACCTCA 622

Qy        250 CCACATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAA 308
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        621 CCACATCCTTGNTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAA 562

Qy        309 CTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTA 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        561 CTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTA 502

Qy        369 TATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCA 428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 TATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCA 442

Qy        429 CCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTT 488
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        441 CCTTGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTT 382

Qy        489 ACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTG 548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 ACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTG 322

Qy        549 CCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTT 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 CCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTT 262

Qy        609 GCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGG 668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGG 202

Qy        669 GGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAA 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAA 142

Qy        729 CCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCC 788
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 CCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAAACTAAGCC 82

Qy        789 CCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCT 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 CCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGGGTGAGTCT 22

Qy        849 GTTCTTGGCGGCCGC 863
              |||||||||||||||
Db         21 GTTCTTGGCGGCCGC 7


RESULT 6
CA306071/c
LOCUS       CA306071                679 bp    mRNA    linear   EST 01-NOV-2002
DEFINITION  UI-H-FT1-bhs-e-13-0-UI.s1 NCI_CGAP_FT1 Homo sapiens cDNA clone
            UI-H-FT1-bhs-e-13-0-UI 3', mRNA sequence.
ACCESSION   CA306071
VERSION     CA306071.1  GI:24469122
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Gary W. Hunninghake, U of I
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FT1-bhs-e-13-0-UI"
                     /tissue_type="Aveolar Macrophage"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FT1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_FT1 is a normalized cDNA library constructed from
                     a pool of 81 RNA samples from Alveolar Macrophages
                     challenged with different treatments. The library was
                     normalized according to Bonaldo, Lennon and Soares, Genome
                     Research, 6:791-806, 1996. First strand cDNA synthesis was
                     primed with an oligo-dT primer containing a Not I site.
                     Double stranded cDNA was ligated to an EcoR I adaptor,
                     digested with Not I, and cloned directionally into
                     pT7T3-Pac vector. The oligonucleotide used to prime the
                     synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     GGCCATGCCG. The tissue was provided by Dr. Gary W.
                     Hunninghake of the University of Iowa.
                     TAG_TISSUE=Human Lung Aveolar Macrophage
                     TAG_LIB=UI-H-FT1
                     TAG_SEQ=GGCCATGCCG"
ORIGIN

  Query Match           12.5%;  Score 637.4;  DB 14;  Length 679;
  Best Local Similarity  99.2%;  Pred. No. 2e-127;
  Matches  660;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy       4165 CACTTGGCCAGTGCCCACCCAGCCCCTCAGGTG-GCACATCTGGAGGGCAGGGGTTGAGG 4223
              |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db        679 CACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGCNNACATCTGGAGGGCAGGGGTTGAGG 620

Qy       4224 GGCCA-CCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGG 4282
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 GGCCACCCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGG 560
```

```
Qy       4283 ATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATG 4342
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 ATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATG 500

Qy       4343 GTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGA 4402
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 GTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGA 440

Qy       4403 GGTGCCCTATAAACCAGAAGAGAAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACC 4462
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 GGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACC 380

Qy       4463 AGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACC 4522
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 AGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACC 320

Qy       4523 TTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTAT 4582
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 TTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTAT 260

Qy       4583 CTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTT 4642
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 CTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTT 200

Qy       4643 TACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGC 4702
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 TACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGC 140

Qy       4703 CCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATT 4762
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 CCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATT 80

Qy       4763 TATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCC 4822
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          79 TATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCC 20

Qy       4823 CAGAA 4827
              || ||
Db          19 CAAAA 15
```

```
RESULT 7
BX092291
LOCUS       BX092291                744 bp    mRNA    linear   EST 23-JAN-2003
DEFINITION  BX092291 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGp998N021747 ;
            IMAGE:713281, mRNA sequence.
ACCESSION   BX092291
VERSION     BX092291.1  GI:27825213
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998N021747.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998N021747 ; IMAGE:713281"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_GCB1"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
```

```
  Query Match             12.5%;  Score 634.4;  DB 13;  Length 744;
  Best Local Similarity   98.1%;  Pred. No. 9.3e-127;
  Matches  683;  Conservative    0;  Mismatches    8;  Indels    5;  Gaps    4;

Qy        133 CAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 192
              || | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 CAAGCTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 108

Qy        193 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 168

Qy        253 CATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 311
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 CATCCTTGCTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 228

Qy        312 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 288

Qy        372 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 348

Qy        432 TGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 491
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 TGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 408

Qy        492 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 468

Qy        552 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCCCCACCCTTGCA 528

Qy        612 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 GGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACCAAGAACGGGGT 588

Qy        672 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 GAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTTCAATCTAACCC 648

Qy        732 CACCCTACACTTCCCACACTCAGGAATCACA-TCCTAGAATATACCCAAAACTAAG-CCC 789
              |||||||||||||||||| |||||||||| |||||||||||||||||||||||| ||| |
Db        649 CACCCTACACTTCCCACACTTAGGAATCACATTCCTAGAATATACCCAAAACTAAGCCCC 708
```

```
Qy        790 CATAAGGCAGCCCG--ACCCTAGTGGTCTAACCCTA 823
              ||||| |||||||||   || ||| ||||||||||||
Db        709 CATAANGCAGCCCGGACCCTTAGNGGTCTAACCCTA 744


RESULT 8
CA748807/c
LOCUS       CA748807                675 bp    mRNA    linear   EST 26-NOV-2002
DEFINITION  UI-H-FT1-bid-i-12-0-UI.s1 NCI_CGAP_FT1 Homo sapiens cDNA clone
            UI-H-FT1-bid-i-12-0-UI 3', mRNA sequence.
ACCESSION   CA748807
VERSION     CA748807.1  GI:25569227
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 675)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Gary W. Hunninghake, U of I
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .675
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FT1-bid-i-12-0-UI"
                     /tissue_type="Aveolar Macrophage"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FT1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_FT1 is a normalized cDNA library constructed from
                     a pool of 81 RNA samples from Alveolar Macrophages
                     challenged with different treatments. The library was
                     normalized according to Bonaldo, Lennon and Soares, Genome
                     Research, 6:791-806, 1996. First strand cDNA synthesis was
                     primed with an oligo-dT primer containing a Not I site.
                     Double stranded cDNA was ligated to an EcoR I adaptor,
                     digested with Not I, and cloned directionally into
                     pT7T3-Pac vector. The oligonucleotide used to prime the
                     synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     GGCCATGCCG. The tissue was provided by Dr. Gary W.
                     Hunninghake of the University of Iowa.
                     TAG_TISSUE=Human Lung Aveolar Macrophage
                     TAG_LIB=UI-H-FT1
                     TAG_SEQ=GGCCATGCCG"
ORIGIN

  Query Match             12.2%;  Score 618.2;  DB 14;  Length 675;
  Best Local Similarity   98.9%;  Pred. No. 3e-123;
  Matches  654;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    3;

Qy       4171 GCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTGGAGGGCAGGGGTTGAGGGGCCACC 4230
              |||||||||||||||||||||| |||||||||||||||||||| |    ||||||||
Db        674 GCCAGTGCCCACCCAGCCCCTCA-GTGGCACATCTGGAGGGCAGGGTTGA--GGGCCACC 618

Qy       4231 ACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTGT 4290
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||

Db        617 ACCACACAT-CCTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTGT 559

Qy       4291 TGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAAT 4350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 TGCTCCCTTATCCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAAT 499

Qy       4351 TCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCT 4410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 TCTACACAAAAAGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCT 439

Qy       4411 ATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAG 4470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 ATAAACCAGAAGAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAG 379

Qy       4471 GAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCG 4530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GAGTCTCCAAAGCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCG 319

Qy       4531 AATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGT 4590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 AATTGTCTTCCAGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGT 259

Qy       4591 TGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGC 4650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 TGCCAGGCTGGGGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGC 199

Qy       4651 AGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCA 4710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 AGTAGCATTTTGGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCA 139

Qy       4711 CTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGAT 4770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 CTCATGACTCTAAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGAT 79

Qy       4771 GATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAACTT 4830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          78 GATATTTATTGCAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAACTT 19

Qy       4831 A 4831
              |
Db          18 A 18


RESULT 9
BF213570
LOCUS       BF213570                691 bp    mRNA    linear   EST 06-NOV-2000
DEFINITION  601845370F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4070560 5',
            mRNA sequence.
ACCESSION   BF213570
VERSION     BF213570.1  GI:11107156
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM914  row: l  column: 17
            High quality sequence stop: 637.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4070560"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 12.0%; Score 612; DB 10; Length 691;
Best Local Similarity 94.5%; Pred. No. 6.7e-122;
Matches 656; Conservative 0; Mismatches 35; Indels 3; Gaps 2;

```
Qy       3727 GGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTT 3786
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTT 60

Qy       3787 CAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGC 3846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGC 120

Qy       3847 TCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACT 3906
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACT 180

Qy       3907 TTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCC 3966
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCC 240

Qy       3967 TCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACC 4026
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCCACC 300

Qy       4027 TCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTA 4086
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTTCTA 360

Qy       4087 CAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGC 4146
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCATGGC 420

Qy       4147 CTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTG 4206
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACATCTG 480

Qy       4207 GAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTTGGCT 4266
             |||||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||
Db        481 GAGGGCAGGGGTTGAGGGGCCACCA-CAAACATGCCTTTCTGGGGTGAAGCCCTTTGG-- 537

Qy       4267 GCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATCCAAT 4326
               ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        538 TGCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATTCCCAAATCACTCTATACATCCAAT 597

Qy       4327 TCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCAGGGT 4386
             |||||||| ||    |  | ||||||||||||||||||||||||||||||        ||
Db        598 TCAGGAAACCAAACATGGTGGAATTCTACACAAAAAGAGATGAGATTAACCAGTGCAGGC 657

Qy       4387 TGGGGTCTGCATTGGAGGTGCCCTATAAACCAGA 4420
             |||||||| | |  ||    | | |  |||
Db        658 TGGGGTCTGATTGGAGGGGCTATAACAGAAGAGA 691
```

RESULT 10
CC138171
LOCUS CC138171 831 bp DNA linear GSS 16-APR-2003
DEFINITION NDL.3C1.T7 Notre Dame Liverpool Aedes aegypti genomic clone NDL.3C1, genomic survey sequence.
ACCESSION CC138171
VERSION CC138171.1 GI:30007226
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes; Stegomyia.
REFERENCE 1 (bases 1 to 831)
AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.
TITLE BAC end sequencing of Aedes aegypti
JOURNAL Unpublished (2003)
COMMENT Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .831
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="liverpool"
/db_xref="taxon:7159"
/clone="NDL.3C1"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pECBAC1; Site_1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 12.0%; Score 611.6; DB 28; Length 831;
Best Local Similarity 85.9%; Pred. No. 8.9e-122;
Matches 718; Conservative 0; Mismatches 104; Indels 14; Gaps 3;

```
Qy       3313 TGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTCACAGCTGAACTA 3372
             | |||  || |||  |||||||||| |||||||| |||||| |||||||||| |||||
Db          1 TTAGCTGGCTGCTTCGCCCTCCTCCGTCCCCATCGCCCTCTCTTTTTCACAGCTTAACTA 60

Qy       3373 CCTG-GGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAGACTC 3431
              |||| |||||||||||||||| || |||||||||| || || ||||||||||||||||
Db         61 CCTGAGGCCCCCCTTTCAACGAACCTGACTTCAACCCACCTCGGCTGGGGGCAGAGACTC 120

Qy       3432 TGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGGCTGA 3491
             ||||||||||||||||  ||||||| ||||||| || ||||| |||||||| |||||||
Db        121 TGCCCAGGGCCACTGTCAACTTGGATGTGTGGCAAAACCTCAGTGACAAACTTCGGCTGA 180

Qy       3492 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGTCAGG 3551
             ||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||| ||||
Db        181 CCCAGAACTACGAGGCCTACAGCCACCTTCTGTGCTACCTGCGTGGCCTCAACCGCCAGG 240

Qy       3552 CTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGCCTGC 3611
             | ||||| || |||||||||||||||||||||||||||||||||||||||||| ||||
Db        241 CCGCCACAGCCGAGCTGCGCCGCAGCCTGGCACACTTCTGCACCAGCCTCCAGGGTCTGC 300

Qy       3612 TGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTGCCTG 3671
             |||||||||||||||| ||||||||||||||||||||||||||||||||||| |||||||
Db        301 TGGGCAGCATTGCGGGTGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCTCTGCCTG 360

Qy       3672 GGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATGGACG 3731
             ||||||| ||||| ||| |||||| |||||||||||| ||||||||||||||||||||| |
Db        361 GGACTGAGCCCACCTGGTCTCCTGACCCTGCCCACAGCGACTTCCTCCAGAAGATGGATG 420
```

```
Qy       3732 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTCAACC 3791
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||
Db        421 ACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCAGCCAAAGACTTCAACC 480

Qy       3792 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCTCATG 3851
             ||||||||||||||||||||||||||||||||||||||||||||| ||||| ||| ||||
Db        481 GGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCGCCTGGAGGCCCATG 540

Qy       3852 GCTTCTGACTTCTGACCTTCTC--------CTCTTCGCTCCCCCTTCAAACCCTGCTCCC 3903
             ||||||||  |||||||||  |         || | | |||| ||||||||||||| |||||
Db        541 GCTTCTGATCTCTGACCTTTACCACCTTCTCTTTCCCCTCCTCCTTCAAACCCTACTCCC 600

Qy       3904 ACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGA 3963
             ||| || ||| |  |  ||||| |||||||||  ||||||||||||||||||||  |||
Db        601 ACTCTGGGAGGGAGAAACCTGTGTGCCAACACTCGTTGAGCCAGGAGACAGAAGCCATGA 660

Qy       3964 GCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCCC 4023
             ||||||||||| || |   ||  | | ||||||||| || || |||||| |||||||
Db        661 GCCTCTGGCCCTCCCCAGACTTGGAGGAGGGTGTGATGCAATAAGGCCTGTCCCCTCCCC 720

Qy       4024 ACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTTT 4083
               | ||||||||  |||| |  ||||||| ||||| || |||||     ||  |||  |||
Db        721 GCTTCCCAAAGTCCTACAGGTCTGGGGAAGAGGTGCAATAGGC-----CCCATCCCATTT 775

Qy       4084 CTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGC 4139
               | |||||||| ||||| |||||   |||||||||| ||||| ||  |||| |
Db        776 CCACAGGAAGTAATGCTCAAGGGAGCCCGAAGTGGTTCAAGTTGGCGCCAAGGCTC 831
```

RESULT 11
CK032835/c
LOCUS CK032835 793 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT_15196860 NIH_MGC_195 Homo sapiens cDNA clone IMAGE:7002176 5', mRNA sequence.
ACCESSION CK032835
VERSION CK032835.1 GI:38558759
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRBK4 row: e column: 06
High quality sequence start: 4
High quality sequence stop: 618.
FEATURES Location/Qualifiers
source 1. .793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002176"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC_195"
/note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2: loxP-HindIII; Clones from this library have been PCR-amplified using gene-specific primers to contain the complete open reading frame (based on known gene sequences available from NCBI's RefSeq). Template for PCR is cDNA derived from either pooled cytoplasmic polyA RNA from 30 cells lines or pooled total RNA from 10 different tissues (from BD Biosciences/Clontech and Washington University). PCR products are directionally cloned into the loxP sites of the pDNR-Dual vector. Library constructed by Dr. Narayan Bhat, Earl Bere III and Hongling Liao (Gene Expression Laboratory, Research Technology Program, SAIC Frederick, NCI-Frederick, Frederick, MD 21702). For information on which gene each clone represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat a Note: this is a NIH_MGC Library."
ORIGIN

Query Match 11.6%; Score 588; DB 14; Length 793;
Best Local Similarity 99.8%; Pred. No. 1.2e-116;
Matches 599; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy       3368 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAG 3427
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 AACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCAGAG 553

Qy       3428 ACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGG 3487
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 ACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTGCGG 493

Qy       3488 CTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGT 3547
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAACCGT 433

Qy       3548 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 3607
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 CAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAGGGC 373

Qy       3608 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 3667
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCGCTG 313

Qy       3668 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 3727
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAGATG 253

Qy       3728 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTC 3787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGACTTC 193

Qy       3788 AACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCT 3847
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 AACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGGGCT 133

Qy       3848 CATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCACTT 3907
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        132 CATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCCCAC-T 74

Qy       3908 TGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCT 3967
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          73 TGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTGAGCCT 14
```

RESULT 12
CA309862/c
LOCUS CA309862 606 bp mRNA linear EST 01-NOV-2002
DEFINITION UI-H-FT1-bid-a-09-0-UI.s1 NCI_CGAP_FT1 Homo sapiens cDNA clone UI-H-FT1-bid-a-09-0-UI 3', mRNA sequence.
ACCESSION CA309862
VERSION CA309862.1 GI:24472916
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 606)

```
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Gary W. Hunninghake, U of I
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .606
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FT1-bid-a-09-0-UI"
                     /tissue_type="Aveolar Macrophage"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FT1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_FT1 is a normalized cDNA library constructed from
                     a pool of 81 RNA samples from Alveolar Macrophages
                     challenged with different treatments. The library was
                     normalized according to Bonaldo, Lennon and Soares, Genome
                     Research, 6:791-806, 1996. First strand cDNA synthesis was
                     primed with an oligo-dT primer containing a Not I site.
                     Double stranded cDNA was ligated to an EcoR I adaptor,
                     digested with Not I, and cloned directionally into
                     pT7T3-Pac vector. The oligonucleotide used to prime the
                     synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     GGCCATGCCG. The tissue was provided by Dr. Gary W.
                     Hunninghake of the University of Iowa.
                     TAG_TISSUE=Human Lung Aveolar Macrophage
                     TAG_LIB=UI-H-FT1
                     TAG_SEQ=GGCCATGCCG"
ORIGIN

  Query Match           11.4%;  Score 578;  DB 14;  Length 606;
  Best Local Similarity   99.8%;  Pred. No. 1.6e-114;
  Matches  589;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       4242 CTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTAT 4301
             ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        606 CTTTCTGGGGTGAAGCCCTTTGGCTGCCCCACTCTCCTTGGATGGGTG-TGCTCCCTTAT 548

Qy       4302 CCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAA 4361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 CCCCAAATCACTCTATACATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAA 488

Qy       4362 AGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAA 4421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 AGAGATGAGATTAACAGTGCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAA 428

Qy       4422 GAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAA 4481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGAAAATACTGAAAGCACAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAA 368

Qy       4482 GCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCC 4541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GCACAGAGTGGCAAACAAAACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCC 308

Qy       4542 AGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGG 4601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 AGTATTACGGTGCCTCTTCTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGG 248

Qy       4602 GGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTT 4661
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GGAGGGCAACCATAGCCACACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTT 188

Qy       4662 GGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCT 4721
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GGGGTGTAGGGTGGCAGCTCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCT 128

Qy       4722 AAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTG 4781
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AAGTGTGTTGTATTAATATTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTG 68

Qy       4782 CAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAACTTA 4831
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          67 CAGAATTTCTATTCTTGTATTAACAAATAAAATGCTTGCCCCAGAACTTA 18


RESULT 13
CF271924/c
LOCUS       CF271924                763 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AGENCOURT_15197020 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002181 5', mRNA sequence.
ACCESSION   CF271924
VERSION     CF271924.2  GI:38453378
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627836.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 11
            High quality sequence start: 7
            High quality sequence stop: 648.
FEATURES             Location/Qualifiers
     source          1. .763
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002181"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
```

ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
a Note: this is a NIH_MGC Library."
ORIGIN

Query Match 11.4%; Score 577.8; DB 14; Length 763;
Best Local Similarity 99.3%; Pred. No. 1.9e-114;
Matches 601; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 557

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 497

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 437

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 377

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 317

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 257

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 197

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        196 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCCGCACCTGG 137

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        136 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 77

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          76 CAC-TTGTGAGAGCCAGCCC-GTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 19

Qy       3963 AGCCT 3967
             |||||
Db          18 AGCCT 14
```

RESULT 14
CF271930/c
LOCUS CF271930 801 bp mRNA linear EST 19-NOV-2003
DEFINITION AGENCOURT_15196924 NIH_MGC_195 Homo sapiens cDNA clone IMAGE:7002178 5', mRNA sequence.
ACCESSION CF271930
VERSION CF271930.2 GI:38453381
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Aug 12, 2003 this sequence version replaced gi:33627842.
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: Bhat Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRBK4 row: e column: 08
High quality sequence start: 6
High quality sequence stop: 646.

FEATURES Location/Qualifiers
source 1. .801
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7002178"
/tissue_type="mixed"
/lab_host="DH5A (T1 phage-resistant)"
/clone_lib="NIH_MGC_195"
/note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2: loxP-HindIII; Clones from this library have been PCR-amplified using gene-specific primers to contain the complete open reading frame (based on known gene sequences available from NCBI's RefSeq). Template for PCR is cDNA derived from either pooled cytoplasmic polyA RNA from 30 cells lines or pooled total RNA from 10 different tissues (from BD Biosciences/Clontech and Washington University). PCR products are directionally cloned into the loxP sites of the pDNR-Dual vector. Library constructed by Dr. Narayan Bhat, Earl Bere III and Hongling Liao (Gene Expression Laboratory, Research Technology Program, SAIC Frederick, NCI-Frederick, Frederick, MD 21702). For information on which gene each clone represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat a Note: this is a NIH_MGC Library."
ORIGIN

Query Match 11.3%; Score 576.8; DB 14; Length 801;
Best Local Similarity 99.2%; Pred. No. 3.2e-114;
Matches 600; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        616 ATCTGAACTACNTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 557

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        556 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 497

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        496 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 437

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        436 ACCGTCGGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 377

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        376 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 317

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        316 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 257

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        256 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 197

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
```

```
Db        196 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 137

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902

Db        136 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 77

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962

Db         76 CAC-TTGTGAGAGCCAGCCC-GTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 19

Qy       3963 AGCCT 3967

Db         18 AGCCT 14


RESULT 15
CF271926/c
LOCUS       CF271926                800 bp    mRNA    linear   EST 19-NOV-2003
DEFINITION  AGENCOURT_15196988 NIH_MGC_195 Homo sapiens cDNA clone
            IMAGE:7002180 5', mRNA sequence.
ACCESSION   CF271926
VERSION     CF271926.2  GI:38453379
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 800)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Aug 12, 2003 this sequence version replaced gi:33627838.
            Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Narayan Bhat
             cDNA Library Preparation: Bhat Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBK4  row: e  column: 10
            High quality sequence start: 8
            High quality sequence stop: 720.
FEATURES             Location/Qualifiers
     source          1. .800
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7002180"
                     /tissue_type="mixed"
                     /lab_host="DH5A (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_195"
                     /note="Vector: pDNR-Dual; Site_1: loxP-SalI; Site_2:
                     loxP-HindIII; Clones from this library have been
                     PCR-amplified using gene-specific primers to contain the
                     complete open reading frame (based on known gene sequences
                     available from NCBI's RefSeq). Template for PCR is cDNA
                     derived from either pooled cytoplasmic polyA RNA from 30
                     cells lines or pooled total RNA from 10 different tissues
                     (from BD Biosciences/Clontech and Washington University).
                     PCR products are directionally cloned into the loxP sites
                     of the pDNR-Dual vector. Library constructed by Dr.
                     Narayan Bhat, Earl Bere III and Hongling Liao (Gene
                     Expression Laboratory, Research Technology Program, SAIC
                     Frederick, NCI-Frederick, Frederick, MD 21702). For
                     information on which gene each clone represents, please
                     visit our anonymous ftp site at
                     ftp://image.llnl.gov/image/rearrayed_plates/IRBK.preSV.dat
                     a Note: this is a NIH_MGC Library."
ORIGIN

  Query Match           11.3%;  Score 572.8;  DB 14;  Length 800;
  Best Local Similarity  99.3%;  Pred. No. 2.4e-113;
  Matches  596;  Conservative    0;  Mismatches    2;  Indels    2;  Gaps    2;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422

Db        614 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 555

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482

Db        554 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 495

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542

Db        494 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 435

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602

Db        434 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 375

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662

Db        374 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 315

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722

Db        314 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 255

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782

Db        254 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 195

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842

Db        194 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 135

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902

Db        134 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 75

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962

Db         74 CACTTTG-GAGAGCCAGCCC-GTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGAG 17


Search completed: August 16, 2004, 04:01:38
Job time : 11585.8 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 12:15:49 ; Search time 1741.71 Seconds
(without alignments)
12407.700 Million cell updates/sec

Title: US-09-931-704-3
Perfect score: 5087
Sequence: 1 aacctgcgagtgggcctggc.....cctttgctaagtcttcctca 5087

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	5087	100.0	5087	3	AAA39482	Aaa39482 Human NNT
	2	5087	100.0	5087	6	ABK11648	Abk11648 Human nov
	3	5076	99.8	5088	2	AAV22653	Aav22653 Human gen
	4	5076	99.8	5088	2	AAV47511	Aav47511 Human neu
	5	1464.4	28.8	1790	3	AAA88546	Aaa88546 Human int
	6	1463.4	28.8	1710	2	AAX16161	Aax16161 Human car
	7	535.4	10.5	1008	4	AAK51548	Aak51548 Human pol
	8	529.4	10.4	729	4	AAD04201	Aad04201 Human car
	9	525.4	10.3	881	4	AAH74484	Aah74484 Nucleotid
	10	523.4	10.3	797	2	AAV22652	Aav22652 cDNA enco
	11	523.4	10.3	797	2	AAV47510	Aav47510 Human neu
	12	523.4	10.3	797	3	AAA39481	Aaa39481 Human NNT
	13	523.4	10.3	797	6	ABK11647	Abk11647 Human cDN
	14	495.4	9.7	968	4	ABA09140	Aba09140 Human car
	15	495.4	9.7	968	4	AAK52532	Aak52532 Human pol
c	16	493.4	9.7	495	4	AAI25564	Aai25564 Probe #15
c	17	493.4	9.7	495	4	ABA71773	Aba71773 Human foe
c	18	493.4	9.7	495	4	AAI52108	Aai52108 Probe #20
c	19	493.4	9.7	495	4	ABA37855	Aba37855 Probe #16
c	20	493.4	9.7	495	4	AAK46202	Aak46202 Human bon
c	21	493.4	9.7	495	4	AAK20142	Aak20142 Human bra
c	22	493.4	9.7	495	4	ABS45921	Abs45921 Human liv
c	23	493.4	9.7	495	6	ABS20513	Abs20513 Human gen
	24	492.4	9.7	768	4	AAH99772	Aah99772 Human pro
c	25	492	9.7	492	4	AAI16384	Aai16384 Probe #63
c	26	492	9.7	492	4	ABA59256	Aba59256 Human foe
c	27	492	9.7	492	4	AAI39047	Aai39047 Probe #77
c	28	492	9.7	492	4	ABA27996	Aba27996 Probe #64
c	29	492	9.7	492	4	AAK33253	Aak33253 Human bon
c	30	492	9.7	492	4	AAK07469	Aak07469 Human bra
c	31	492	9.7	492	4	ABS33003	Abs33003 Human liv
c	32	492	9.7	492	6	ABS08085	Abs08085 Human gen
	33	437.2	8.6	465	6	ABL81689	Abl81689 Human ova
	34	429.6	8.4	819	2	AAV22654	Aav22654 cDNA enco
	35	429.6	8.4	819	2	AAV47512	Aav47512 Mouse neu
	36	429.6	8.4	819	3	AAA39483	Aaa39483 Murine NN
	37	429.6	8.4	819	6	ABK11649	Abk11649 Mouse cDN
c	38	429.4	8.4	432	6	ABL81658	Abl81658 Human ova
	39	421.8	8.3	648	3	AAA88547	Aaa88547 Mouse int
	40	398	7.8	457	8	ACH22723	Ach22723 Human adu
c	41	355.8	7.0	360	6	ABL79632	Abl79632 Human ova
	42	330	6.5	342	2	AAX51546	Aax51546 Human sec
	43	329	6.5	340	3	AAC14792	Aac14792 Human sec
	44	324.6	6.4	396	2	AAX16162	Aax16162 Human car
	45	250.6	4.9	269	6	ABL82404	Abl82404 Human ova

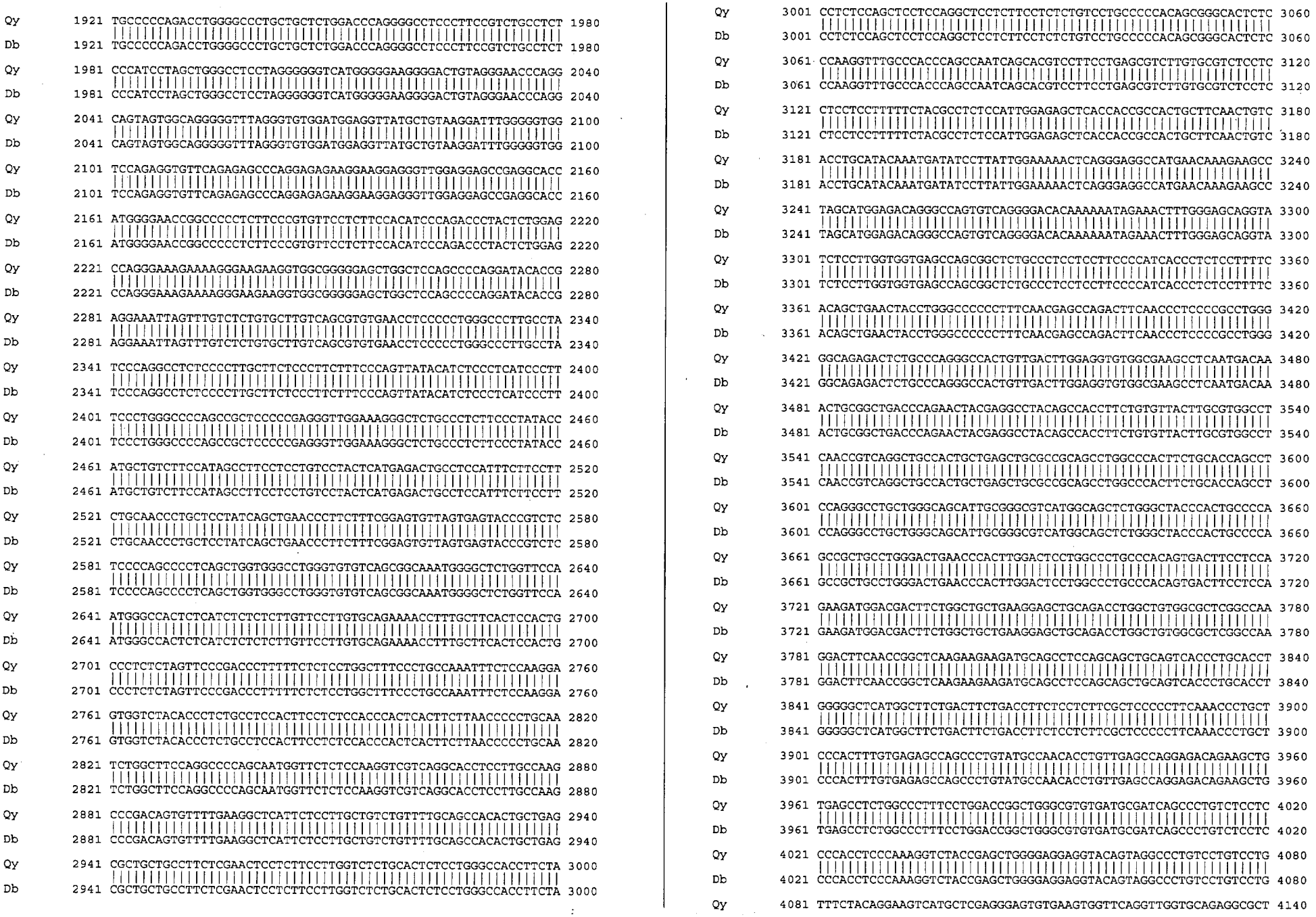
ALIGNMENTS

RESULT 1
AAA39482
ID AAA39482 standard; DNA; 5087 BP.
XX
AC AAA39482;
XX
DT 24-AUG-2000 (first entry)
XX
DE Human NNT-1 DNA.
XX
KW NNT-1; human; neurotrophic factor; nootropic; neuroprotective; treatment;
KW anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW retinopathy; immune disorder; hematopoietic disorder; ss.
XX
OS Homo sapiens.
XX
PN US6054294-A.
XX
PD 25-APR-2000.
XX
PF 12-DEC-1997; 97US-00988819.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 2000-338492/29.
XX
PT New nucleic acids encoding neurotrophic factors useful for stimulating
PT growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS Claim 1b; Col 31-36; 42pp; English.
XX
CC This invention describes a novel nucleic acid molecule (I) encoding a
CC novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC which are useful for treating patients in whom various cells of the
CC central, autonomic, or peripheral nervous system have degenerated and/or
CC have been damaged by congenital disease, trauma, mechanical damage,
CC surgery, stroke, ischemia, infection, metabolic disease, nutritional

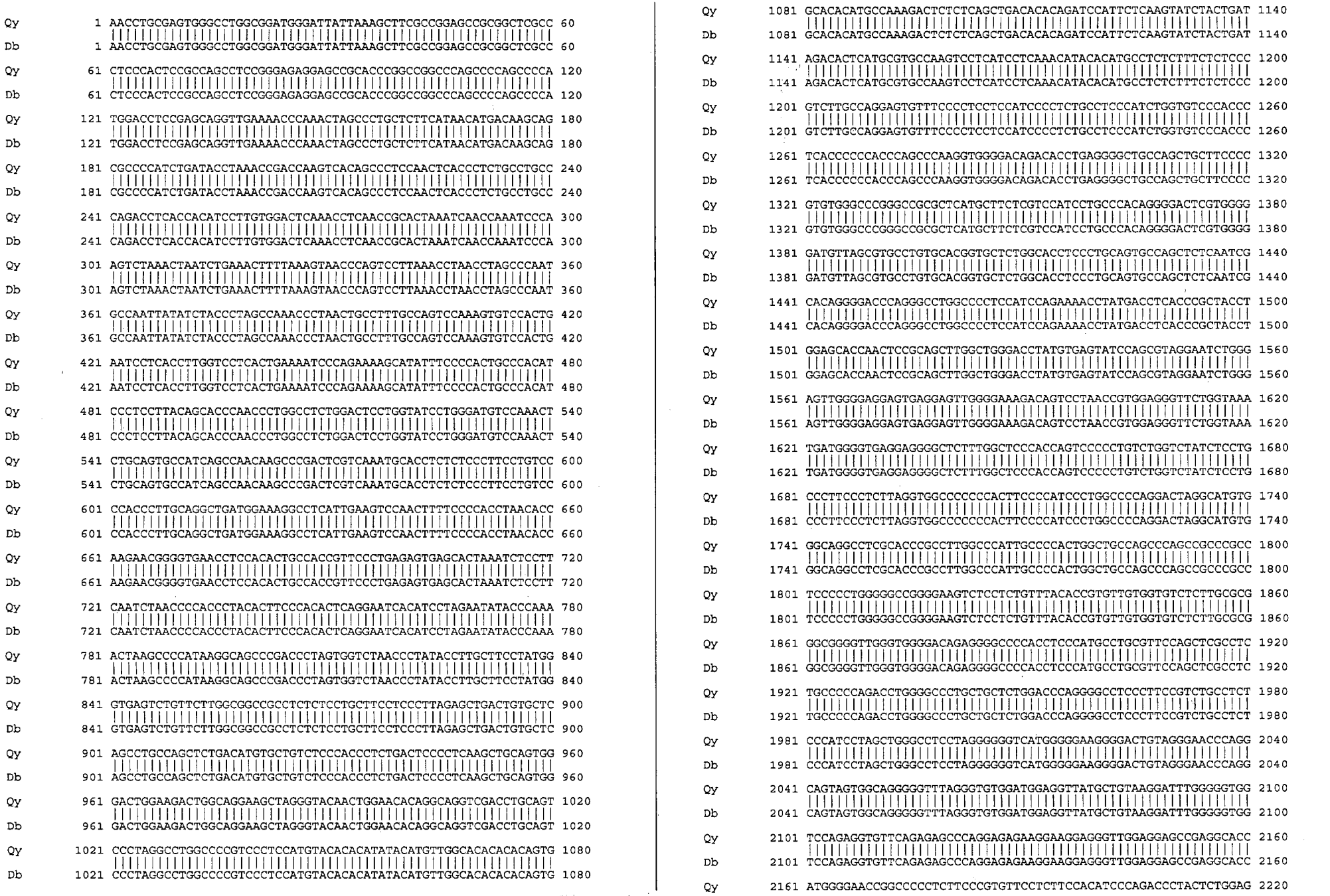
CC deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC neuropathy induced by diabetes or other metabolic disorders, and/or
CC dystrophies or degeneration of the neural retina such as retinitis
CC pigmentosa, drug-induced retinopathies, stationary forms of night
CC blindness, progressive cone-rod degeneration, immune disorders and
CC hematopoietic disorders. (I) is effective in treating neurological
CC conditions and promotes neuron regeneration. Neural functions are
CC effectively restored in patients suffering from various neurological
CC disorders. This sequence encodes the human NNT-1 protein described in the
CC method of the invention
XX
SQ Sequence 5087 BP; 992 A; 1746 C; 1191 G; 1158 T; 0 U; 0 Other;

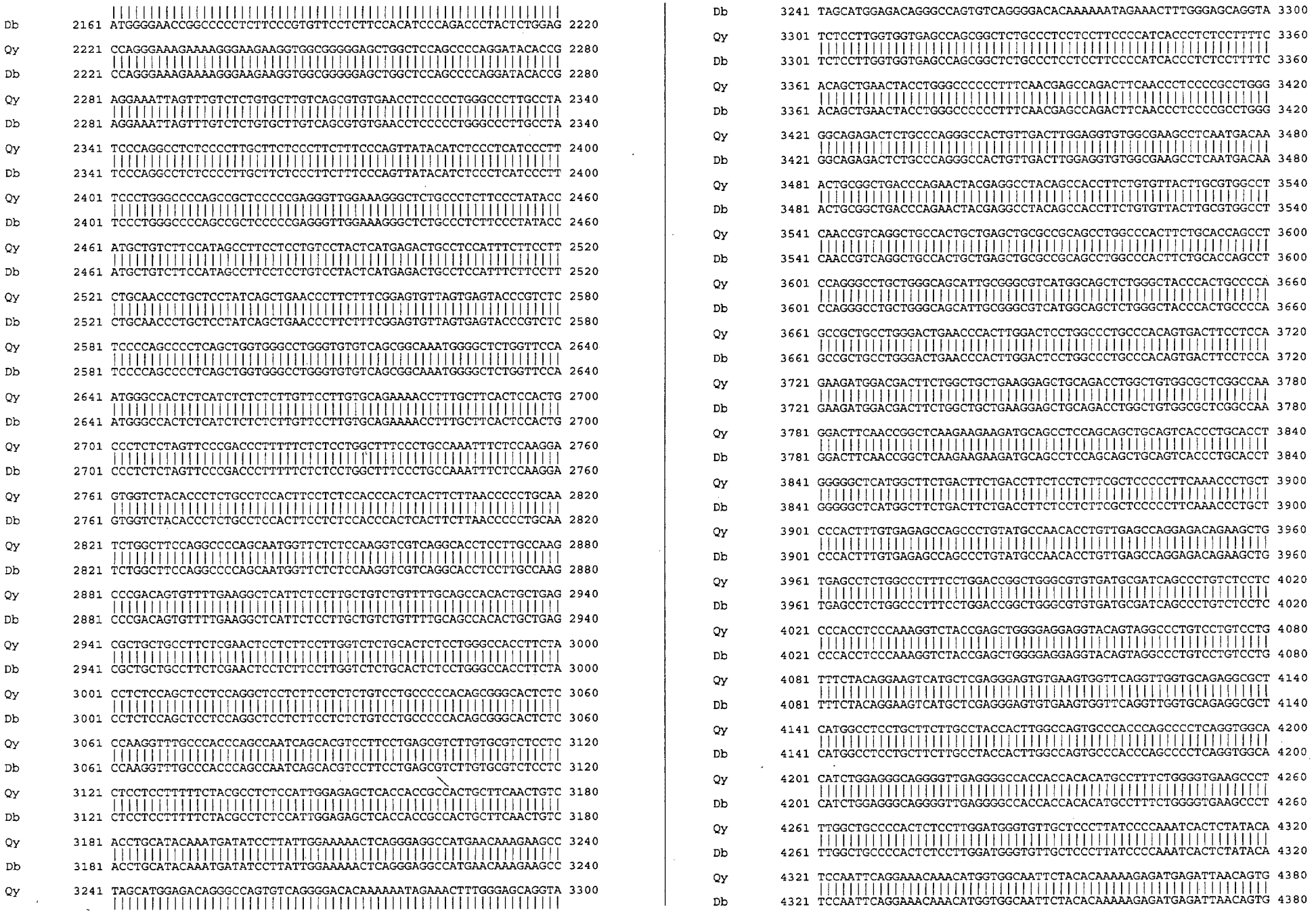
Query Match 100.0%; Score 5087; DB 3; Length 5087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780

Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900

Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960

Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020

Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080

Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140

Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200

Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260

Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320

Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380

Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440

Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500

Qy       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560

Qy       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620

Qy       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680

Qy       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740

Qy       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800

Qy       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860

Qy       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
```


Qy 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Qy 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Db 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy 4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140

Qy       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200

Qy       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260

Qy       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320

Qy       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380

Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440

Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500

Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
                |||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087


RESULT 2
ABK11648
ID   ABK11648 standard; DNA; 5087 BP.
XX
AC   ABK11648;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human novel neurotrophic factor NNT1, genomic DNA.
XX
KW   Human; ds; gene; NNT1; neurotrophic factor; IgE-related disease;
KW   Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW   pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW   vascular restenosis; rheumatoid arthritis; psoriatic arthritis;
KW   inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW   autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW   inflammatory bowel disease; transplant rejection; reproductive disorder;
KW   graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    137. .138
FT                   /*tag= a
FT                   /note= "Intervening, unsequenced region of more than 1KB"
XX
PN   WO200215977-A2.
XX
PD   28-FEB-2002.
XX
PF   17-AUG-2001; 2001WO-US025906.
XX
PR   18-AUG-2000; 2000US-0226436P.
PR   16-AUG-2001; 2001US-00931704.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Senaldi G;
XX
DR   WPI; 2002-280867/32.
XX
PT   Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT   patient, preventing IgE-related disease and treating allergic diseases,
PT   involves administering NNT-1 inhibitor to a patient.
XX
PS   Claim 2; Fig 2; 63pp; English.
XX
CC   The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC   modulating IgE levels in a patient, preventing an IgE-related disease,
CC   and treating allergic diseases, comprising administering a
CC   therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC   inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC   related disease or susceptibility to an IgE-related disease, by
CC   determining the presence or amount of expression of an NNT1 polypeptide
CC   encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC   occurring variant, and diagnosing an IgE-related disease or
CC   susceptibility of an IgE-related disease based on the presence or amount
CC   of expression of the polypeptide and a pharmaceutical composition for use
CC   in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC   inhibitor is useful for preventing and treating IgE-related disease,
CC   modulating IgE levels, and treating allergic diseases e.g. Type I
CC   allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC   asthma, immune diseases and disorders, diseases involving abnormal cell
CC   proliferation including cancer, arteriosclerosis and vascular restenosis,
CC   diseases and conditions relating to dysfunction of immune system
CC   including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC   arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC   disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC   transplant rejection, and graft versus host disease, and reproductive
CC   diseases and disorders including infertility, miscarriage, preterm labour
CC   and delivery, and endometriosis. The present sequence is the genomic DNA
CC   for human NNT1
XX
SQ   Sequence 5087 BP; 992 A; 1746 C; 1191 G; 1158 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 5087;  DB 6;  Length 5087;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```





```
Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440

Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500

Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 3
AAV22653
ID AAV22653 standard; DNA; 5088 BP.
XX
AC AAV22653;
XX
DT 13-JUL-1998 (first entry)
XX
DE Human genomic DNA encoding neurotrophic factor NNT-1.
XX
KW Human; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW treatment; neurological disease; degeneration; Parkinson's disease;
KW amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 138
FT /*tag= a
FT /note= "represents intervening unsequenced region of 1
FT Kb"
XX
PN US5741772-A.
XX
PD 21-APR-1998.
XX
PF 03-FEB-1997; 97US-00792019.
XX
PR 03-FEB-1997; 97US-00792019.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M;
XX
DR WPI; 1998-260526/23.
XX
PT Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT for stimulating growth of motor and sympathetic neurons.
XX
PS Disclosure; Fig 2; 41pp; English.
XX
CC The present sequence encodes a human neurotrophic factor, designated NNT-
CC 1, which is capable of stimulating growth of motor or sympathetic
CC neurons. The NNT-1 protein is useful in the treatment of neurological
CC diseases characterised by the degeneration and death of particular
CC classes of neurons. These diseases specifically include Parkinson's
CC disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke
CC and various degenerative disorders affecting vision
XX
SQ Sequence 5088 BP; 992 A; 1746 C; 1191 G; 1158 T; 0 U; 1 Other;

Query Match 99.8%; Score 5076; DB 2; Length 5088;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGT-TGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCA 179
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTNTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCA 180

Qy        180 GCGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGC 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGC 240

Qy        240 CCAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCC 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCC 300

Qy        300 AAGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAA 360

Qy        360 TGCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACT 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACT 420

Qy        420 GAATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACA 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACA 480

Qy        480 TCCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAAC 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAAC 540

Qy        540 TCTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTC 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTC 600

Qy        600 CCCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACAC 659
```

```
Db        601 CCCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACAC 660
Qy        660 CAAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCT 719
Db        661 CAAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCT 720
Qy        720 TCAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAA 779
Db        721 TCAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAA 780
Qy        780 AACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATG 839
Db        781 AACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATG 840
Qy        840 GGTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCT 899
Db        841 GGTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCT 900
Qy        900 CAGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTG 959
Db        901 CAGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTG 960
Qy        960 GGACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAG 1019
Db        961 GGACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAG 1020
Qy       1020 TCCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGT 1079
Db       1021 TCCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGT 1080
Qy       1080 GGCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGA 1139
Db       1081 GGCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGA 1140
Qy       1140 TAGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCC 1199
Db       1141 TAGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCC 1200
Qy       1200 CGTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACC 1259
Db       1201 CGTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACC 1260
Qy       1260 CTCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCC 1319
Db       1261 CTCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCC 1320
Qy       1320 CGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGG 1379
Db       1321 CGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGG 1380
Qy       1380 GGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATC 1439
Db       1381 GGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATC 1440
Qy       1440 GCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACC 1499
Db       1441 GCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACC 1500
Qy       1500 TGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGG 1559
Db       1501 TGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGG 1560
Qy       1560 GAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAA 1619
Db       1561 GAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAA 1620
Qy       1620 ATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCT 1679
Db       1621 ATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCT 1680
Qy       1680 GCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGT 1739
Db       1681 GCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGT 1740
Qy       1740 GGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGC 1799
Db       1741 GGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGC 1800
Qy       1800 CTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGC 1859
Db       1801 CTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGC 1860
Qy       1860 GGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCT 1919
Db       1861 GGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCT 1920
Qy       1920 CTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTC 1979
Db       1921 CTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTC 1980
Qy       1980 TCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAG 2039
Db       1981 TCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAG 2040
Qy       2040 GCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTG 2099
Db       2041 GCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTG 2100
Qy       2100 GTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCAC 2159
Db       2101 GTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCAC 2160
Qy       2160 CATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGA 2219
Db       2161 CATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGA 2220
Qy       2220 GCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACC 2279
Db       2221 GCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACC 2280
Qy       2280 GAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCT 2339
Db       2281 GAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCT 2340
Qy       2340 ATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCT 2399
Db       2341 ATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCT 2400
Qy       2400 TTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATAC 2459
Db       2401 TTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATAC 2460
Qy       2460 CATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCT 2519
Db       2461 CATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCT 2520
Qy       2520 TCTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCT 2579
Db       2521 TCTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCT 2580
Qy       2580 CTCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCC 2639
Db       2581 CTCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCC 2640
Qy       2640 AATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACT 2699
Db       2641 AATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACT 2700
Qy       2700 GCCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGG 2759
Db       2701 GCCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGG 2760
Qy       2760 AGTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCA 2819
Db       2761 AGTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCA 2820
```

```
Qy       2820 ATCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAA 2879
Db       2821 ATCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAA 2880

Qy       2880 GCCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGA 2939
Db       2881 GCCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGA 2940

Qy       2940 GCGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCT 2999
Db       2941 GCGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCT 3000

Qy       3000 ACCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCT 3059
Db       3001 ACCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCT 3060

Qy       3060 CCCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCT 3119
Db       3061 CCCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCT 3120

Qy       3120 CCTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGT 3179
Db       3121 CCTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGT 3180

Qy       3180 CACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGC 3239
Db       3181 CACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGC 3240

Qy       3240 CTAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGT 3299
Db       3241 CTAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGT 3300

Qy       3300 ATCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTT 3359
Db       3301 ATCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTT 3360

Qy       3360 CACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 3419
Db       3361 CACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 3420

Qy       3420 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 3479
Db       3421 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 3480

Qy       3480 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 3539
Db       3481 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 3540

Qy       3540 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 3599
Db       3541 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 3600

Qy       3600 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 3659
Db       3601 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 3660

Qy       3660 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 3719
Db       3661 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 3720

Qy       3720 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 3779
Db       3721 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 3780

Qy       3780 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 3839
Db       3781 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 3840

Qy       3840 TGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 3899
Db       3841 TGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 3900

Qy       3900 TCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCT 3959
Db       3901 TCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCT 3960

Qy       3960 GTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCT 4019
Db       3961 GTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCT 4020

Qy       4020 CCCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCT 4079
Db       4021 CCCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCT 4080

Qy       4080 GTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGC 4139
Db       4081 GTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGC 4140

Qy       4140 TCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGC 4199
Db       4141 TCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGC 4200

Qy       4200 ACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCC 4259
Db       4201 ACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCC 4260

Qy       4260 TTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATAC 4319
Db       4261 TTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATAC 4320

Qy       4320 ATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGT 4379
Db       4321 ATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGT 4380

Qy       4380 GCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCA 4439
Db       4381 GCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCA 4440

Qy       4440 CAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAA 4499
Db       4441 CAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAA 4500

Qy       4500 AACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTT 4559
Db       4501 AACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTT 4560

Qy       4560 CTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCA 4619
Db       4561 CTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCA 4620

Qy       4620 CACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGC 4679
Db       4621 CACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGC 4680

Qy       4680 TCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATA 4739
Db       4681 TCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATA 4740

Qy       4740 TTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGT 4799
Db       4741 TTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGT 4800

Qy       4800 ATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTG 4859
Db       4801 ATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTG 4860

Qy       4860 GTGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTG 4919
Db       4861 GTGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTG 4920

Qy       4920 CACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAG 4979
Db       4921 CACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAG 4980

Qy       4980 ACTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTT 5039
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 ACTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTT 5040

Qy       5040 CTGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 CTGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5088


RESULT 4
AAV47511
ID   AAV47511 standard; cDNA; 5088 BP.
XX
AC   AAV47511;
XX
DT   09-NOV-1998  (first entry)
XX
DE   Human neurotrophic factor NNT-1 genomic DNA.
XX
KW   NNT-1; neurotrophic factor; human; antiinflammatory; adjuvant;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   peripheral neurophathy; dystrophy; neural retina degeneration;
KW   common variable immunodeficiency; CVID; selective IgA deficiency;
KW   hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW   therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1. .119
FT                   /*tag=  a
FT   CDS             120. .3857
FT                   /*tag=  b
FT                   /note= "contains 2 introns"
FT   exon            120. .135
FT                   /*tag=  c
FT                   /number= 1
FT   intron          136. .1368
FT                   /*tag=  d
FT                   /number= 1
FT   misc_feature    137. .138
FT                   /*tag=  e
FT                   /note= "intervening unsequenced region of >1kb"
FT   exon            1369. .1535
FT                   /*tag=  f
FT                   /number= 2
FT   intron          1536. .3365
FT                   /*tag=  g
FT                   /number= 2
FT   exon            3366. .3857
FT                   /*tag=  h
FT                   /number= 3
XX
PN   WO9833922-A1.
XX
PD   06-AUG-1998.
XX
PF   02-FEB-1998;   98WO-US002363.
XX
PR   03-FEB-1997;   97US-00792019.
PR   30-JAN-1998;   98US-00016534.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M,  Elliot GS,  Senaldi G,  Sarmiento U;
XX
DR   WPI; 1998-437475/37.
DR   P-PSDB; AAW29715.
XX
PT   Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT   NNT-1 - useful for treatment of neurological and immunological diseases
PT   or inflammation, also as vaccine adjuvant.
XX
PS   Claim 4; Fig 2; 120pp; English.
XX
CC   This newly isolated human genomic DNA sequence (deposited at ATCC 98294)
CC   codes for a novel neurotrophic factor, designated NNT-1 (see AAW29715),
CC   that is a growth factor for neurons and for B or T cells. It was obtained
CC   from a human genomic P1 library using human NNT-1 cDNA (see AAV47510) as
CC   probe. Vectors containing the cDNA or genomic DNA and host cells are
CC   provided for use in the production of NNT-1 polypeptides. These are used
CC   to treat: (i) neurological or immunological diseases, specifically
CC   Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic lateral
CC   sclerosis, Charcot-Marie-Tooth syndrome, peripheral neuropathy, dystrophy
CC   and degeneration of the neural retina, or conditions characterised by T
CC   or B cell defects, e.g. common variable immunodeficiency (CVID),
CC   selective IgA deficiency, hypogammaglobulinaemia and X-linked
CC   agammaglobulinaemia (claimed), but many others disclosed; and (ii)
CC   inflammation. NNT-1 is also able to boost immunoreactivity and antibody
CC   production following vaccination, and, since it inhibits tumour necrosis
CC   factor production, it may also be useful for treating sepsis. NNT-1
CC   nucleic acid fragments are also used as hybridisation probes in
CC   diagnostic assays. In addition, cells that have been engineered to
CC   express NNT-1 can be implanted, or nucleic acids are delivered in gene
CC   therapy vectors
XX
SQ   Sequence 5088 BP; 992 A; 1746 C; 1191 G; 1158 T; 0 U; 1 Other;

  Query Match             99.8%;  Score 5076;  DB 2;  Length 5088;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCC 299
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGSTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCC 300

Qy        300 AAGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAA 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAA 360

Qy        360 TGCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACT 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACT 420

Qy        420 GAATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACA 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACA 480

Qy        480 TCCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAAC 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAAC 540

Qy        540 TCTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTC 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTC 600

Qy        600 CCCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACAC 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACAC 660
```

```
Qy       660 CAAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCT 719
Db       661 CAAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCT 720

Qy       720 TCAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAA 779
Db       721 TCAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAA 780

Qy       780 AACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATG 839
Db       781 AACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATG 840

Qy       840 GGTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCT 899
Db       841 GGTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCT 900

Qy       900 CAGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTG 959
Db       901 CAGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTG 960

Qy       960 GGACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAG 1019
Db       961 GGACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAG 1020

Qy      1020 TCCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGT 1079
Db      1021 TCCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGT 1080

Qy      1080 GGCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGA 1139
Db      1081 GGCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGA 1140

Qy      1140 TAGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCC 1199
Db      1141 TAGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCC 1200

Qy      1200 CGTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACC 1259
Db      1201 CGTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACC 1260

Qy      1260 CTCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCC 1319
Db      1261 CTCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCC 1320

Qy      1320 CGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGG 1379
Db      1321 CGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGG 1380

Qy      1380 GGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATC 1439
Db      1381 GGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATC 1440

Qy      1440 GCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACC 1499
Db      1441 GCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACC 1500

Qy      1500 TGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGG 1559
Db      1501 TGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGG 1560

Qy      1560 GAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAA 1619
Db      1561 GAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAA 1620

Qy      1620 ATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCT 1679
Db      1621 ATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCT 1680

Qy      1680 GCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGT 1739
Db      1681 GCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGT 1740

Qy      1740 GGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGC 1799
Db      1741 GGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGC 1800

Qy      1800 CTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGC 1859
Db      1801 CTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGC 1860

Qy      1860 GGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCT 1919
Db      1861 GGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCT 1920

Qy      1920 CTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTC 1979
Db      1921 CTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTC 1980

Qy      1980 TCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAG 2039
Db      1981 TCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAG 2040

Qy      2040 GCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTG 2099
Db      2041 GCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTG 2100

Qy      2100 GTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCAC 2159
Db      2101 GTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCAC 2160

Qy      2160 CATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGA 2219
Db      2161 CATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGA 2220

Qy      2220 GCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACC 2279
Db      2221 GCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACC 2280

Qy      2280 GAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCT 2339
Db      2281 GAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCT 2340

Qy      2340 ATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCT 2399
Db      2341 ATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCT 2400

Qy      2400 TTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATAC 2459
Db      2401 TTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATAC 2460

Qy      2460 CATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCT 2519
Db      2461 CATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCT 2520

Qy      2520 TCTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCT 2579
Db      2521 TCTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCT 2580

Qy      2580 CTCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCC 2639
Db      2581 CTCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCC 2640

Qy      2640 AATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACT 2699
Db      2641 AATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACT 2700

Qy      2700 GCCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGG 2759
Db      2701 GCCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGG 2760

Qy      2760 AGTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCA 2819
Db      2761 AGTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCA 2820

Qy      2820 ATCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAA 2879
```

```
Db      2821 ATCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAA 2880
Qy      2880 GCCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGA 2939
Db      2881 GCCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGA 2940
Qy      2940 GCGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCT 2999
Db      2941 GCGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCT 3000
Qy      3000 ACCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCT 3059
Db      3001 ACCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCT 3060
Qy      3060 CCCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCT 3119
Db      3061 CCCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCT 3120
Qy      3120 CCTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGT 3179
Db      3121 CCTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGT 3180
Qy      3180 CACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGC 3239
Db      3181 CACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGC 3240
Qy      3240 CTAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGT 3299
Db      3241 CTAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGT 3300
Qy      3300 ATCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTT 3359
Db      3301 ATCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTT 3360
Qy      3360 CACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 3419
Db      3361 CACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGG 3420
Qy      3420 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 3479
Db      3421 GGGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACA 3480
Qy      3480 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 3539
Db      3481 AACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCC 3540
Qy      3540 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 3599
Db      3541 TCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCC 3600
Qy      3600 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 3659
Db      3601 TCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCC 3660
Qy      3660 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 3719
Db      3661 AGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCC 3720
Qy      3720 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 3779
Db      3721 AGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCA 3780
Qy      3780 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 3839
Db      3781 AGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACC 3840
Qy      3840 TGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 3899
Db      3841 TGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 3900
Qy      3900 TCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCT 3959
Db      3901 TCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCT 3960
```

```
Qy      3960 GTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCT 4019
Db      3961 GTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCT 4020
Qy      4020 CCCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCT 4079
Db      4021 CCCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCT 4080
Qy      4080 GTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGC 4139
Db      4081 GTTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGC 4140
Qy      4140 TCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGC 4199
Db      4141 TCATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGC 4200
Qy      4200 ACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCC 4259
Db      4201 ACATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCC 4260
Qy      4260 TTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATAC 4319
Db      4261 TTTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATAC 4320
Qy      4320 ATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGT 4379
Db      4321 ATCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGT 4380
Qy      4380 GCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAAATACTGAAAGCA 4439
Db      4381 GCAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAAATACTGAAAGCA 4440
Qy      4440 CAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAA 4499
Db      4441 CAGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAA 4500
Qy      4500 AACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTT 4559
Db      4501 AACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTT 4560
Qy      4560 CTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCA 4619
Db      4561 CTCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCA 4620
Qy      4620 CACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGC 4679
Db      4621 CACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGC 4680
Qy      4680 TCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATA 4739
Db      4681 TCCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATA 4740
Qy      4740 TTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGT 4799
Db      4741 TTTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGT 4800
Qy      4800 ATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTG 4859
Db      4801 ATTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTG 4860
Qy      4860 GTGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTG 4919
Db      4861 GTGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTG 4920
Qy      4920 CACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAG 4979
Db      4921 CACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAG 4980
Qy      4980 ACTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTT 5039
Db      4981 ACTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTT 5040
```

```
Qy       5040 CTGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 CTGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5088
```

RESULT 5
AAA88546
ID AAA88546 standard; DNA; 1790 BP.
XX
AC AAA88546;
XX
DT 22-JAN-2001 (first entry)
XX
DE Human interleukin-B60 (IL-B60) gene.
XX
KW Interleukin-B60; IL-B60; human; cytokine; chromosome 11;
KW cytokine-like factor-1; haematopoietic; inflammation; antiinflammatory;
KW autoimmune disease; therapy; ds.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	CDS	162. .809
FT		/*tag= a
FT	sig_peptide	162. .212
FT		/*tag= b
FT	mat_peptide	213. .806
FT		/*tag= c

XX
PN WO200053631-A1.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US006182.
XX
PR 11-MAR-1999; 99US-00267901.
XX
PA (SCHE) SCHERING CORP.
XX
PI Oppmann B, Timans JC, Kastelein RA, Bazan JF;
XX
DR WPI; 2000-587426/55.
DR P-PSDB; AAB19686.
XX
PT Cytokine-like factor 1 (CLF-1) and interleukin (IL)-B60 complexes,
PT polypeptides, and nucleic acids, useful in research, diagnosis and for
PT treating inflammatory and autoimmune disorders.
XX
PS Claim 17; Page 15-16; 97pp; English.
XX
CC The present sequence is that of DNA encoding human interleukin-B60 (IL-
CC B60, see AAB19586), a novel, small soluble cytokine-like protein of 198
CC amino acids that exhibits structural motifs characteristic of a member of
CC the long-chain cytokines, and which shows homology to granulocyte colony
CC stimulating factor and interleukin-6. IL-60B may have either stimulatory
CC or inhibitory effects on haematopoietic cells, including e.g. lymphoid
CC cells, such as T-cells, B-cells, natural killer cells, macrophages,
CC dendritic cells, haematopoietic progenitors, etc. Methods are provided
CC for modulating the physiology or development of a cell or tissue culture
CC cells by contacting the cell with an agonist or antagonist of IL-B60 or
CC an agonist of antagonist of a complex of mature IL-B60 and its partner,
CC cytokine-like factor-1 (CLF-1, see AAB19588). The IL-B60/CLF-1 cytokine
CC serves as a key physiological factor in motor neuron development and
CC regeneration. IL-60B, its agonists and antagonists may be used to treat
CC inflammatory or autoimmune disorders and also for drug screening. The
CC IL60B gene maps to human chromosome 11
XX
SQ Sequence 1790 BP; 381 A; 560 C; 474 G; 375 T; 0 U; 0 Other;

Query Match 28.8%; Score 1464.4; DB 3; Length 1790;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1465; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 372

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 432

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 492

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 552

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 612

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 672

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 732

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        733 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 792

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        793 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 852

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        853 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 912

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        913 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 972

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        973 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 1032

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1033 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1092

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1152

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1212

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1272

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1332

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1392

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1393 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1452

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1453 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1512

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1513 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1572

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1573 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1632

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1633 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1692

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1693 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1752

Qy       4803 AACAAATAAAATGCTTGCCCCAGAAC 4828
                |||||||||||||||||||||||||
Db       1753 AACAAATAAAATGCTTGCCCCAGAAC 1778
```

RESULT 6
AAX16161
ID AAX16161 standard; DNA; 1710 BP.
XX
AC AAX16161;
XX
DT 28-APR-1999 (first entry)
XX
DE Human cardiotrophin-like cytokine encoding DNA.
XX
KW Human; cardiotrophin-like cytokine; interleukin 6 cytokine family; CLC;
KW IL-6; diagnosis; detection; immune system-related disorder; cancer;
KW cardiac disorder; heart failure; hypertension; cancer;
KW autoimmune disorder; infection; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 46. .723
FT /*tag= a
FT sig_peptide 46. .129
FT /*tag= b
FT mat_peptide 130. .720
FT /*tag= c
FT /product= "cardiotrophin-like cytokine"
XX
PN WO9900415-A1.
XX
PD 07-JAN-1999.
XX
PF 29-JUN-1998; 98WO-US013129.
XX
PR 30-JUN-1997; 97US-0051311P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Shi Y, Ruben SM;
XX
DR WPI; 1999-095678/08.
DR P-PSDB; AAW94466.
XX
PT New isolated cardiotrophin-like cytokine nucleic acid - used to develop
PT products for treating cardiac and immune system disorders, e.g. heart
PT failure, hypertension, cancers, autoimmune disorders and infections.
XX
PS Claim 2; Fig 1; 103pp; English.
XX
CC The present invention relates to a novel cardiotrophin-like cytokine
CC (CLC) protein which is a member of the interleukin 6 (IL-6) cytokine
CC family. The present sequence encodes the human CLC protein. The present
CC invention also describes screening methods for identifying agonists and
CC antagonists of CLC activity, as well as methods for detecting cardiac and
CC immune system-related disorders and therapeutic methods for treating
CC cardiac and immune system-related disorders, e.g heart failure,
CC hypertension, cancers, autoimmune disorders and infections
XX
SQ Sequence 1710 BP; 370 A; 530 C; 448 G; 362 T; 0 U; 0 Other;

Query Match 28.8%; Score 1463.4; DB 2; Length 1710;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 286

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 346

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 406

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 466

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 526

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 586

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 646

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 706

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 766

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 826

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 886

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 946

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1006

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1066
```

```
Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1126

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1186

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1246

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1306

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1366

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1426

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1486

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1546

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1606

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1666

Qy       4803 AACAAATAAAATGCTTGCCCCAGAA 4827
             ||||||||||||||||||||||||||
Db       1667 AACAAATAAAATGCTTGCCCCAGAA 1691
```

RESULT 7
AAK51548
ID AAK51548 standard; cDNA; 1008 BP.
XX
AC AAK51548;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 93.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78415.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 711-712; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 1008 BP; 183 A; 374 C; 247 G; 204 T; 0 U; 0 Other;

Query Match 10.5%; Score 535.4; DB 4; Length 1008;
Best Local Similarity 99.8%; Pred. No. 3.4e-122;
Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 531

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 591

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 651

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 711

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 771

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 831

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 891

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 951

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 3899
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGC 1008
```

```
RESULT 8
AAD04201
ID   AAD04201 standard; cDNA; 729 BP.
XX
AC   AAD04201;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human cardiotrophin-like cytokine (CLC) cDNA.
XX
KW   Human; biologically active complex; haemopoietin receptor; NR6;
KW   cardiotrophin-like cytokine; CLC; therapy; prophylaxis; proliferation;
KW   differentiation; cell survival; neurotrophic activity; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             11. .679
FT                   /*tag=  a
FT                   /product= "Human cardiotrophin-like cytokine (CLC)
FT                   protein"
FT                   /note= "CDS does not include stop codon"
FT                   /partial
FT   sig_peptide     11. .91
FT                   /*tag=  b
FT   mat_peptide     92. .679
FT                   /*tag=  c
FT                   /product= "Human mature cardiotrophin-like cytokine (CLC)
FT                   protein"
XX
PN   WO200127157-A1.
XX
PD   19-APR-2001.
XX
PF   06-OCT-2000; 2000WO-AU001216.
XX
PR   08-OCT-1999;   99AU-00003327.
PR   12-MAY-2000; 2000AU-00007489.
XX
PA   (AMRA-) AMRAD OPERATIONS PTY LTD.
XX
PI   Nash A,  Jachno KM,  Fabri LJ,  Reid K,  Bartlett PF,  Hilton DJ;
PI   Nakata Y,  Hasegawa M;
XX
DR   WPI; 2001-281978/29.
DR   P-PSDB; AAE00828.
XX
PT   New biologically active complex comprising NR6 and cardiotrophin-like-
PT   cytokine, for facilitating proliferation, differentiation and/or survival
PT   of a cell.
XX
PS   Claim 31; Page 112-114; 123pp; English.
XX
CC   The present invention relates to a biologically active complex comprising
CC   a haemopoietin receptor, NR6 and cardiotrophin-like cytokine (CLC). The
CC   complex is useful in the manufacture of-a medicament for the treatment
CC   and/or prophylaxis of a subject, as it is involved in facilitating
CC   proliferation, differentiation and/or survival of a cell. The complex or
CC   its components have neurotrophic activity. The present sequence is human
CC   cardiotrophin-like cytokine (CLC) cDNA
XX
SQ   Sequence 729 BP; 132 A; 261 C; 196 G; 140 T; 0 U; 0 Other;

  Query Match            10.4%;  Score 529.4;  DB 4;  Length 729;
  Best Local Similarity   99.8%;  Pred. No. 8.9e-121;
  Matches  530;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 251

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 311

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 371

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 431

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 491

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 551

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 611

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 671

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAA 3893
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAA 722


RESULT 9
AAH74484
ID   AAH74484 standard; DNA; 881 BP.
XX
AC   AAH74484;
XX
DT   15-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of a human NNT-1 protein.
XX
KW   NNT-1; CLF-1; sCNTFRalpha; nervous system; neuron; nervous system;
KW   neuro-muscular function; tumour; immune system; haematopoietic system;
KW   reproductive system; liver; skeletal muscle; neurodegenerative disease;
KW   amyotrophic lateral sclerosis; Parkinson's disease; Huntingdon's disease;
KW   muscular mass; paralysis; cancer; obesity; fertility; endometriosis;
KW   blastocyst implantation; thrombosis; retinal disease;
KW   retinal pigmentosis; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             174. .851
FT                   /*tag=  a
FT                   /product= "NNT-1"
XX
PN   WO200155172-A2.
XX
PD   02-AUG-2001.
XX
PF   26-JAN-2001; 2001WO-FR000253.
XX
PR   27-JAN-2000; 2000FR-00001035.
PR   12-OCT-2000; 2000FR-00013089.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
PA   (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Elson G,  Gauchat J,  Plun-Favreau H,  Chevalier S,  Gascan H;
XX
DR   WPI; 2001-488773/53.
```

```
DR   P-PSDB; AAG63543.
XX
PT   A complex comprising a NNT-1 protein and a CLF-1 and/or sCNTFRalpha
PT   protein useful to treat neurodegenerative disease including Parkinson's
PT   and Huntington's, obesity and cancer.
XX
PS   Disclosure; Page 57-58; 67pp; French.
XX
CC   The present sequence encodes a human NNT-1 protein. The specification
CC   describes a complex comprising a NNT-1 protein and a CLF-1 and/or
CC   sCNTFRalpha protein. The NNT-1/CLF-1 complex is used to modulate activity
CC   of the sCNTFRalpha/gp130/LIFRbeta receptor complex, or to induce
CC   phosphorylation of the tyrosine of gp130 and LIFRbeta, particularly where
CC   cells expressing the receptor complex are in the central or peripheral
CC   nervous system, in neurons implicated in neuro-muscular function or in
CC   skeletal muscle. The complex or antibodies are also used to decrease the
CC   survival, growth or proliferation of tumour cells or to facilitate the
CC   proliferation and/or inhibit differentiation of cells stocks. The complex
CC   is also used to modulate activity of the gp130/LIFRbeta receptor or cells
CC   expressing that receptor, particularly those cells implicated in the
CC   immune, haematopoietic, nervous or reproductive system, the liver or
CC   skeletal muscle. Molecules of the invention may be used to prevent or
CC   treat neurodegenerative diseases including amyotrophic lateral sclerosis,
CC   Parkinson's and Huntington's disease, to repair or regenerate nervous or
CC   muscular tissue or to maintain muscular mass in paralysis patients. They
CC   may also be used to treat cancer, obesity and associated diseases, and to
CC   improve fertility, particularly to avoid endometriosis and/or assist
CC   blastocyst implantation, thrombosis, or retinal disease, particular
CC   retinal pigmentosis
XX
SQ   Sequence 881 BP; 158 A; 318 C; 246 G; 159 T; 0 U; 0 Other;

  Query Match            10.3%;  Score 525.4;  DB 4;  Length 881;
  Best Local Similarity  99.8%;  Pred. No. 9.6e-120;
  Matches  526;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 414

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 474

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 534

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 594

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 654

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 714

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 774

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 834

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 3889
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTT 881
```

```
RESULT 10
AAV22652
ID   AAV22652 standard; cDNA; 797 BP.
XX
AC   AAV22652;
XX
DT   13-JUL-1998  (first entry)
XX
DE   cDNA encoding human neurotrophic factor NNT-1.
XX
KW   Human; neurotrophic factor; NNT-1; growth; motor; sympathetic; neuron;
KW   treatment; neurological disease; degeneration; Parkinson's disease;
KW   amyotropic lateral sclerosis; ALS; Alzheimers's disease; stroke; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             90. .767
FT                   /*tag=  a
FT   sig_peptide     90. .170
FT                   /*tag=  b
FT   mat_peptide     171. .764
FT                   /*tag=  c
XX
PN   US5741772-A.
XX
PD   21-APR-1998.
XX
PF   03-FEB-1997;   97US-00792019.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 1998-260526/23.
DR   P-PSDB; AAW56141.
XX
PT   Neurotrophic factor NNT-1 polypeptide and related nucleic acids - useful
PT   for stimulating growth of motor and sympathetic neurons.
XX
PS   Disclosure; Fig 1; 41pp; English.
XX
CC   The present sequence encodes a human neurotrophic factor, designated NNT-
CC   1, which is capable of stimulating growth of motor or sympathetic
CC   neurons. The NNT-1 protein is useful in the treatment of neurological
CC   diseases characterised by the degeneration and death of particular
CC   classes of neurons. These diseases specifically include Parkinson's
CC   disease, amyotropic lateral sclerosis (ALS), Alzheimers's disease, stroke
CC   and various degenerative disorders affecting vision
XX
SQ   Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

  Query Match            10.3%;  Score 523.4;  DB 2;  Length 797;
  Best Local Similarity  99.8%;  Pred. No. 2.9e-119;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510
```

```
Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
             |||||||||||||||||||||||||||||||||||||||||||||
Db        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
```

RESULT 11
AAV47510
ID AAV47510 standard; cDNA; 797 BP.
XX
AC AAV47510;
XX
DT 09-NOV-1998 (first entry)
XX
DE Human neurotrophic factor NNT-1 cDNA.
XX
KW NNT-1; neurotrophic factor; human; antiinflammatory; adjuvant;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW peripheral neurophathy; dystrophy; neural retina degeneration;
KW common variable immunodeficiency; CVID; selective IgA deficiency;
KW hypogammaglobulinaemia; X-linked agammaglobulinaemia; antiseptic;
KW therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 90. .767
FT /*tag= a
FT sig_peptide 90. .170
FT /*tag= b
FT mat_peptide 171. .764
FT /*tag= c
XX
PN WO9833922-A1.
XX
PD 06-AUG-1998.
XX
PF 02-FEB-1998; 98WO-US002363.
XX
PR 03-FEB-1997; 97US-00792019.
PR 30-JAN-1998; 98US-00016534.
XX
PA (AMGE-) AMGEN INC.
XX
PI Chang M, Elliot GS, Senaldi G, Sarmiento U;
XX
DR WPI; 1998-437475/37.
DR P-PSDB; AAW29715.
XX
PT Newly isolated nucleic acid encoding human or murine neurotrophic factor
PT NNT-1 - useful for treatment of neurological and immunological diseases
PT or inflammation, also as vaccine adjuvant.
XX
PS Claim 3; Fig 1; 120pp; English.
XX
CC This newly isolated human cDNA sequence (deposited at ATCC 98295) codes
CC for a novel neurotrophic factor, designated NNT-1 (see AAW29715), that is
CC a growth factor for neurons and for B or T cells. It was obtained from a
CC T-cell lymphoma cDNA library by expressed sequence tag analysis on the
CC basis of homology to CNTF. The isolated NNT-1 cDNA was used as probe to
CC isolated NNT-1 genomic DNA (see AAV47511). Vectors containing the cDNA or
CC genomic DNA and host cells are provided for use in the production of NNT-
CC 1 polypeptides. These are used to treat: (i) neurological or
CC immunological diseases, specifically Alzheimer's, Parkinson's or
CC Huntington's diseases, amyotrophic lateral sclerosis, Charcot-Marie-Tooth
CC syndrome, peripheral neuropathy, dystrophy and degeneration of the neural
CC retina, or conditions characterised by T or B cell defects, e.g. common
CC variable immunodeficiency (CVID), selective IgA deficiency,
CC hypogammaglobulinaemia and X-linked agammaglobulinaemia (claimed), but
CC many others disclosed; and (ii) inflammation. NNT-1 is also able to boost
CC immunoreactivity and antibody production following vaccination, and,
CC since it inhibits tumour necrosis factor production, it may also be
CC useful for treating sepsis. NNT-1 nucleic acid fragments are also used as
CC hybridisation probes in diagnostic assays. In addition, cells that have
CC been engineered to express NNT-1 can be implanted, or nucleic acids are
CC delivered in gene therapy vectors
XX
SQ Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

Query Match 10.3%; Score 523.4; DB 2; Length 797;
Best Local Similarity 99.8%; Pred. No. 2.9e-119;
Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
             |||||||||||||||||||||||||||||||||||||||||||||
Db        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
```

RESULT 12
AAA39481
ID AAA39481 standard; cDNA; 797 BP.
XX
AC AAA39481;
XX
DT 24-AUG-2000 (first entry)
XX
DE Human NNT-1 cDNA.

```
XX
KW   NNT-1; human; neurotrophic factor; nootropic; neuroprotective; treatment;
KW   anticonvulsant; antiparkinsonian; antidiabetic; ophthalmological;
KW   nervous system degeneration; Alzheimer's disease; Parkinson's disease;
KW   amyotrophic lateral sclerosis; Charcot-Marie-Tooth syndrome;
KW   Huntington's disease; peripheral neuropathy; neural retina degeneration;
KW   retinopathy; immune disorder; hematopoietic disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             90. .767
FT                   /*tag=  a
FT                   /product= "NNT-1"
XX
PN   US6054294-A.
XX
PD   25-APR-2000.
XX
PF   12-DEC-1997;   97US-00988819.
XX
PR   03-FEB-1997;   97US-00792019.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Chang M;
XX
DR   WPI; 2000-338492/29.
DR   P-PSDB; AAY87813.
XX
PT   New nucleic acids encoding neurotrophic factors useful for stimulating
PT   growth of motor or sympathetic neurons for treating neuron cell damage.
XX
PS   Claim 1a; Fig 1; 42pp; English.
XX
CC   This invention describes a novel nucleic acid molecule (I) encoding a
CC   novel neurotrophic factor (NNT-1) (II) which has nootropic,
CC   neuroprotective, anticonvulsant, antiparkinsonian, antidiabetic and
CC   ophthalmological activity. (I) is useful for producing NNT-1 polypeptides
CC   which are useful for treating patients in whom various cells of the
CC   central, autonomic, or peripheral nervous system have degenerated and/or
CC   have been damaged by congenital disease, trauma, mechanical damage,
CC   surgery, stroke, ischemia, infection, metabolic disease, nutritional
CC   deficiency, malignancy and/or toxic agents. NNT-1 proteins are used to
CC   treat diseases like Alzheimer's, Parkinson's, amyotrophic lateral
CC   sclerosis, Charcot-Marie-Tooth syndrome, Huntington's disease, peripheral
CC   neuropathy induced by diabetes or other metabolic disorders, and/or
CC   dystrophies or degeneration of the neural retina such as retinitis
CC   pigmentosa, drug-induced retinopathies, stationary forms of night
CC   blindness, progressive cone-rod degeneration, immune disorders and
CC   hematopoietic disorders. (I) is effective in treating neurological
CC   conditions and promotes neuron regeneration. Neural functions are
CC   effectively restored in patients suffering from various neurological
CC   disorders. This sequence encodes the human NNT-1 protein described in the
CC   method of the invention
XX
SQ   Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

  Query Match             10.3%;  Score 523.4;  DB 3;  Length 797;
  Best Local Similarity   99.8%;  Pred. No. 2.9e-119;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
             |||||||||||||||||||||||||||||||||||||||||||||
Db        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
```

```
RESULT 13
ABK11647
ID   ABK11647 standard; cDNA; 797 BP.
XX
AC   ABK11647;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Human cDNA encoding novel neurotrophic factor NNT1.
XX
KW   Human; ss; gene; NNT1; neurotrophic factor; IgE-related disease;
KW   Type I allergic disease; allergic rhinitis; eczema; dermatitis;
KW   pollinosis; asthma; immune disease; cancer; arteriosclerosis;
KW   vascular restenosis; rheumatoid arthritis; psoriatic arthritis;
KW   inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW   autoimmune disease; multiple sclerosis; lupus; diabetes; endometriosis;
KW   inflammatory bowel disease; transplant rejection; reproductive disorder;
KW   graft versus host disease; infertility; miscarriage; preterm labour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             90. .767
FT                   /*tag=  a
FT                   /product= "NNT1"
XX
PN   WO200215977-A2.
XX
PD   28-FEB-2002.
XX
PF   17-AUG-2001; 2001WO-US025906.
XX
PR   18-AUG-2000; 2000US-0226436P.
PR   16-AUG-2001; 2001US-00931704.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Senaldi G;
XX
DR   WPI; 2002-280867/32.
DR   P-PSDB; AAU78176.
XX
PT   Treating Immunoglobulin E-related disease, modulating IgE levels in a
PT   patient, preventing IgE-related disease and treating allergic diseases,
PT   involves administering NNT-1 inhibitor to a patient.
XX
PS   Claim 2; Fig 1; 63pp; English.
```

XX
CC The invention relates to treating Immunoglobulin E (IgE)-related disease,
CC modulating IgE levels in a patient, preventing an IgE-related disease,
CC and treating allergic diseases, comprising administering a
CC therapeutically effective amount of novel neurotrophic factor (NNT)-1
CC inhibitor to a patient. Also included are a method of diagnosing an IgE-
CC related disease or susceptibility to an IgE-related disease, by
CC determining the presence or amount of expression of an NNT1 polypeptide
CC encoded by a NNT1 nucleotide sequence, its fragment or naturally
CC occurring variant, and diagnosing an IgE-related disease or
CC susceptibility of an IgE-related disease based on the presence or amount
CC of expression of the polypeptide and a pharmaceutical composition for use
CC in treating IgE-related disease, comprising the NNT1 inhibitor. The NNT1
CC inhibitor is useful for preventing and treating IgE-related disease,
CC modulating IgE levels, and treating allergic diseases e.g. Type I
CC allergic disease, allergic rhinitis, eczema, dermatitis, pollinosis,
CC asthma, immune diseases and disorders, diseases involving abnormal cell
CC proliferation including cancer, arteriosclerosis and vascular restenosis,
CC diseases and conditions relating to dysfunction of immune system
CC including rheumatoid arthritis, psoriatic arthritis, inflammatory
CC arthritis, osteoarthritis, inflammatory joint disease, autoimmune
CC disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
CC transplant rejection, and graft versus host disease, and reproductive
CC diseases and disorders including infertility, miscarriage, preterm labour
CC and delivery, and endometriosis. The present sequence encodes human NNT1
XX
SQ Sequence 797 BP; 139 A; 297 C; 218 G; 143 T; 0 U; 0 Other;

Query Match 10.3%; Score 523.4; DB 6; Length 797;
Best Local Similarity 99.8%; Pred. No. 2.9e-119;
Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
             |||||||||||||||||||||||||||||||||||||||||||||
Db        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
```

RESULT 14
ABA09140
ID ABA09140 standard; cDNA; 968 BP.
XX

AC ABA09140;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human cardiotrophin-like cytokine homologue cDNA, SEQ ID NO:916.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
DR P-PSDB; ABB11896.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 1; Page 793-794; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.

```
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention
XX
SQ   Sequence 968 BP; 179 A; 353 C; 244 G; 192 T; 0 U; 0 Other;

  Query Match             9.7%;  Score 495.4;  DB 4;  Length 968;
  Best Local Similarity   99.8%;  Pred. No. 2.8e-112;
  Matches  496;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 531

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 591

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 651

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 711

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 771

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 831

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 891

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 951

Qy       3843 GGGCTCATGGCTTCTGA 3859
             |||||||||||||||||
Db        952 GGGCTCATGGCTTCTGA 968


RESULT 15
AAK52532
ID   AAK52532 standard; cDNA; 968 BP.
XX
AC   AAK52532;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 2061.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW   vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW   tissue growth factor; immunomodulatory; cancer; leukaemia;
KW   nervous system disorder; arthritis; inflammation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157190-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US004098.
XX
PR   03-FEB-2000; 2000US-00496914.
PR   27-APR-2000; 2000US-00560875.
PR   20-JUN-2000; 2000US-00598075.
PR   19-JUL-2000; 2000US-00620325.
PR   01-SEP-2000; 2000US-00654936.
PR   15-SEP-2000; 2000US-00663561.
PR   20-OCT-2000; 2000US-00693325.
PR   30-NOV-2000; 2000US-00728422.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT,  Asundi V,  Zhou P,  Xu C,  Cao Y;
PI   Ma Y,  Zhao QA,  Wang D,  Wang J,  Zhang J,  Ren F,  Chen R,  Wang ZW;
PI   Xue AJ,  Yang Y,  Wejhrman T,  Goodrich R;
XX
DR   WPI; 2001-476283/51.
DR   P-PSDB; AAM79399.
XX
PT   Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT   in diagnosis and gene therapy.
XX
PS   Claim 1; Page 4448-4448; 6221pp; English.
XX
CC   The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC   encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC   cytokine, cell proliferation or cell differentiation or which may induce
CC   production of other cytokines in other cell populations. The
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 968 BP; 179 A; 353 C; 244 G; 192 T; 0 U; 0 Other;

  Query Match             9.7%;  Score 495.4;  DB 4;  Length 968;
  Best Local Similarity   99.8%;  Pred. No. 2.8e-112;
  Matches  496;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 531

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 591

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 651

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 711

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 771

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 831

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 891
```

```
Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 951

Qy       3843 GGGCTCATGGCTTCTGA 3859
              |||||||||||||||||
Db        952 GGGCTCATGGCTTCTGA 968
```

Search completed: August 15, 2004, 16:49:08
Job time : 1748.71 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 16:49:33 ; Search time 2941.55 Seconds
(without alignments)
8485.292 Million cell updates/sec

Title: US-09-931-704-3
Perfect score: 5087
Sequence: 1 aacctgcgagtgggcctggc.....cctttgctaagtcttcctca 5087

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

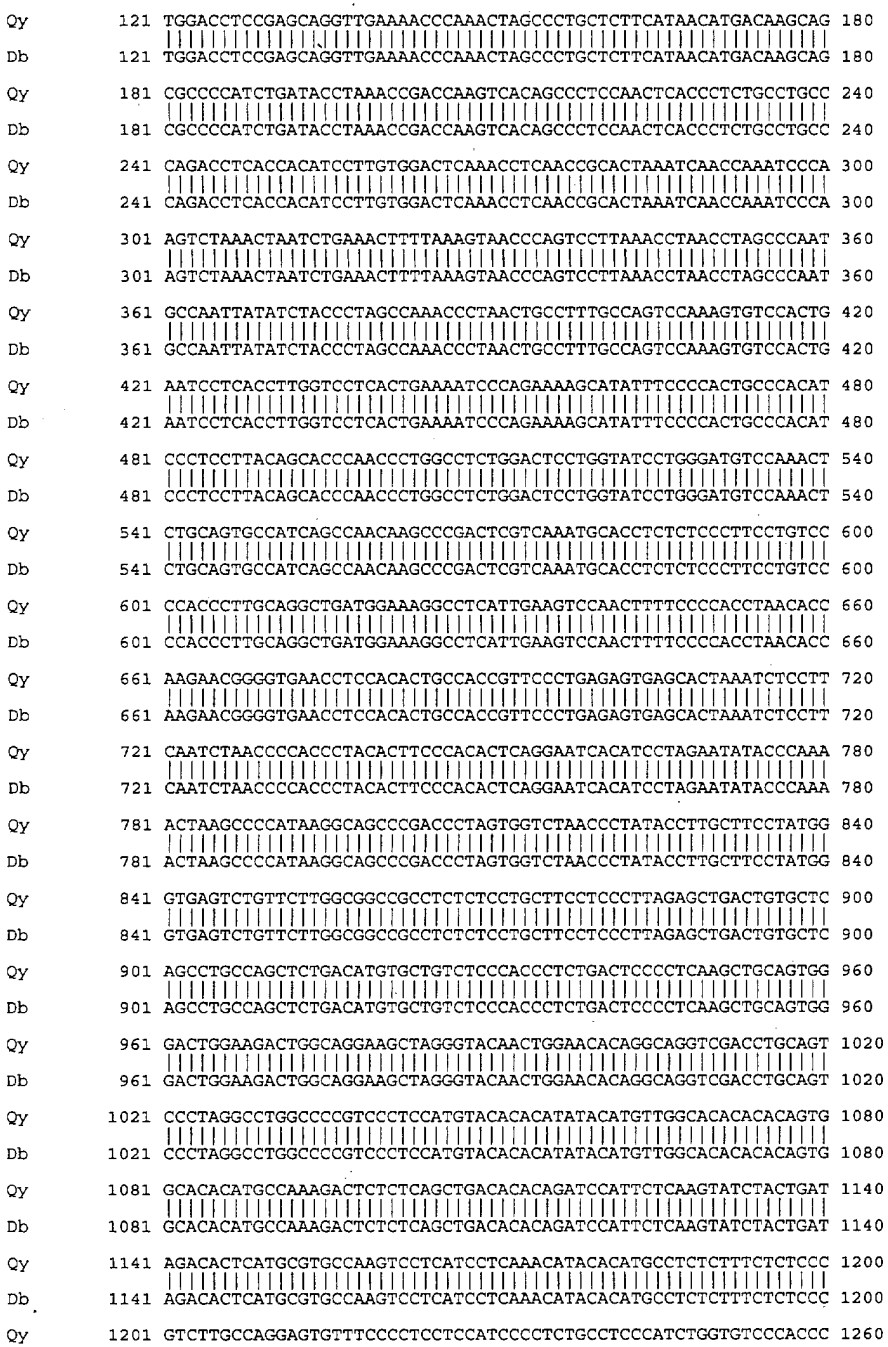
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	5087	100.0	5087	9	US-09-931-704-3	Sequence 3, Appli
	2	1463.4	28.8	1710	15	US-10-212-793-1	Sequence 1, Appli
	3	735.4	14.5	809	13	US-10-027-632-134702	Sequence 134702,
	4	735.4	14.5	809	16	US-10-027-632-134702	Sequence 134702,
	5	630.6	12.4	631	13	US-10-027-632-134703	Sequence 134703,
	6	630.6	12.4	631	13	US-10-027-632-134704	Sequence 134704,
	7	630.6	12.4	631	16	US-10-027-632-134703	Sequence 134703,
	8	630.6	12.4	631	16	US-10-027-632-134704	Sequence 134704,
	9	523.4	10.3	797	9	US-09-931-704-1	Sequence 1, Appli
	10	495.4	9.7	968	13	US-10-276-774-916	Sequence 916, App
c	11	493.4	9.7	495	9	US-09-864-761-23175	Sequence 23175, A
	12	492.4	9.7	768	13	US-10-296-115-607	Sequence 607, App
c	13	492	9.7	492	9	US-09-864-761-6462	Sequence 6462, Ap
	14	437.2	8.6	465	9	US-09-867-701-4667	Sequence 4667, Ap
	15	429.6	8.4	819	9	US-09-931-704-4	Sequence 4, Appli
c	16	429.4	8.4	432	9	US-09-867-701-4636	Sequence 4636, Ap
	17	398	7.8	457	10	US-09-918-995-9935	Sequence 9935, Ap
c	18	355.8	7.0	360	9	US-09-867-701-2610	Sequence 2610, Ap
	19	324.6	6.4	396	15	US-10-212-793-7	Sequence 7, Appli
	20	250.6	4.9	269	9	US-09-867-701-5382	Sequence 5382, Ap
	21	192.8	3.8	231	9	US-09-867-701-2639	Sequence 2639, Ap
	22	165.2	3.2	283	9	US-09-765-231A-75	Sequence 75, Appl
	23	70.8	1.4	2750	13	US-09-854-867-362	Sequence 362, App
	24	66.8	1.3	12050	17	US-10-300-820-4	Sequence 4, Appli
	25	60	1.2	60	10	US-09-908-975-14602	Sequence 14602, A
	26	60	1.2	60	10	US-09-908-975-32089	Sequence 32089, A
	27	60	1.2	60	10	US-09-908-975-32301	Sequence 32301, A
c	28	59.2	1.2	26451	13	US-10-087-192-487	Sequence 487, App
	29	58.6	1.2	65	10	US-09-908-975-29446	Sequence 29446, A
c	30	58.6	1.2	28055	16	US-10-085-117-94	Sequence 94, Appl
	31	57.8	1.1	2778	17	US-10-437-963-56837	Sequence 56837, A
c	32	57.6	1.1	5452	15	US-10-017-161-1481	Sequence 1481, Ap
c	33	57.6	1.1	5452	16	US-10-292-798-1189	Sequence 1189, Ap
	34	57.6	1.1	53522	10	US-09-904-968A-1	Sequence 1, Appli
	35	56	1.1	186391	13	US-10-087-192-136	Sequence 136, App
	36	54.4	1.1	1064	9	US-09-804-682-29	Sequence 29, Appl
	37	54.2	1.1	574	13	US-10-027-632-139721	Sequence 139721,
	38	54.2	1.1	574	16	US-10-027-632-139721	Sequence 139721,
c	39	54	1.1	671	15	US-10-184-644-346	Sequence 346, App
c	40	54	1.1	671	15	US-10-184-634-346	Sequence 346, App
	41	54	1.1	1239	17	US-10-332-859-215	Sequence 215, App
c	42	53.8	1.1	594	13	US-10-142-426-10	Sequence 10, Appl
c	43	53.8	1.1	594	15	US-10-123-155-10	Sequence 10, Appl
c	44	53.8	1.1	594	15	US-10-146-731-10	Sequence 10, Appl
c	45	53.8	1.1	594	15	US-10-140-472-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-931-704-3
; Sequence 3, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
; APPLICANT: Senaldi, Giorgio
; TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: A-695
; CURRENT APPLICATION NUMBER: US/09/931,704
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/226,436
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 5087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (137)..(138)
; OTHER INFORMATION: product = "INTERVENING UNSEQUENCED REGION OF >1KB"
US-09-931-704-3

Query Match 100.0%; Score 5087; DB 9; Length 5087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
```


Qy 121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Db 121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Qy 181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Db 181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Qy 241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
Db 241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
Qy 301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
Db 301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
Qy 361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
Db 361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
Qy 421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
Db 421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
Qy 481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
Db 481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
Qy 541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
Db 541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
Qy 601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
Db 601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
Qy 661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
Db 661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
Qy 721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
Db 721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
Qy 781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
Db 781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
Qy 841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Db 841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Qy 901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Db 901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Qy 961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Db 961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Qy 1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Db 1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Qy 1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Db 1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Qy 1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Db 1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Qy 1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Db 1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Qy 1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Db 1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Qy 1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Db 1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Qy 1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Db 1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Qy 1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Db 1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Qy 1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Db 1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Qy 1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Db 1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Qy 1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Db 1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Qy 1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Db 1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Qy 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Qy 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Qy 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Qy 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Qy 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340

```
Db      2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy      2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db      2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy      2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db      2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy      3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db      3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy      3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db      3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy      3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db      3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy      3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db      3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy      3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Db      3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy      3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db      3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy      4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db      4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy      4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Db      4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy      4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Db      4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy      4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Db      4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Qy      4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Db      4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Qy      4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Db      4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Qy      4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Db      4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Qy      4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Db      4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
```

```
Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 2
US-10-212-793-1
; Sequence 1, Application US/10212793
; Publication No. US20030087395A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Cardiotrophin-Like Cytokine
; FILE REFERENCE: PF385D1C1
; CURRENT APPLICATION NUMBER: US/10/212,793
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 09/438,299
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/106,182
; PRIOR FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: US 60/051,311
; PRIOR FILING DATE: 1997-06-30
; NUMBER OF SEQ ID NOS: 24
; SEQ ID NO 1
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; LOCATION: CDS
; OTHER INFORMATION: (46)..(720)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (46)..(126)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (127)..(720)
US-10-212-793-1

Query Match 28.8%; Score 1463.4; DB 15; Length 1710;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 286

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 346

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 406

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 466

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 526

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 586

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 646

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 706

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 766

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 826

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 886

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 946

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1006

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1066

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1126

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1186

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1246

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1306

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1366

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1426

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1486

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1546

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1606

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1666

Qy       4803 AACAAATAAAATGCTTGCCCCAGAA 4827
                 ||||||||||||||||||||||||
Db       1667 AACAAATAAAATGCTTGCCCCAGAA 1691
```

RESULT 3
US-10-027-632-134702
; Sequence 134702, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134702
; LENGTH: 809
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134702

Query Match 14.5%; Score 735.4; DB 13; Length 809;
Best Local Similarity 97.3%; Pred. No. 4.3e-205;
Matches 789; Conservative 1; Mismatches 17; Indels 4; Gaps 4;

```
Qy       1312 TGCTTCCCCGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGA 1371
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTTCCCCGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGA 60
```

```
Qy       1372 CTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGC 1431
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGC 120

Qy       1432 TCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCAC 1491
                 ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||
Db         121 TCTCAATCGCACAGGGGACCCAGGGCCWGGCCCCTCCATCCAGAAAACCTATGACCTCAC 180

Qy       1492 CCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAG 1551
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAG 240

Qy       1552 GAATCTGGGAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGT 1611
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GAATCTGGGAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGT 300

Qy       1612 TCTGGTAAATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTC 1671
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TCTGGTAAATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTC 360

Qy       1672 TATCTCCTGCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACT 1731
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TATCTCCTGCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACT 420

Qy       1732 AGGCATGTGGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAG 1791
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AGGCATGTGGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAG 480

Qy       1792 CCGCCCGCCTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTC 1851
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CCGCCCGCCTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTC 540

Qy       1852 TCTTGCGCGGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCA 1911
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TCTTGCGCGGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCA 600

Qy       1912 GCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCG 1971
                 |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         601 GCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCA-GGGCCTCCCTTCCG 659

Qy       1972 TCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGG 2031
                 |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         660 TCTGCCTCTCCCATCCTAGCTGGGCCTCCTA-GGGGGTCATGGGGGAAGGGGACTGTAGG 718

Qy       2032 GAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATT 2091
                 ||||||| |||||||||||||||| |||||||||| ||||| | |||||||| | | |
Db         719 GAACCCA-GCAGTAGTGGCAGGGGG-TTAGGGTGTGAATGGAAGGTATGCTGTAGGATTG 776

Qy       2092 TGGGGGTGGTCCAGAGGTGTTCAGAGAGCCC 2122
                 ||||| || | | ||| ||| | | || ||
Db         777 GGGGGGGGGCCAAAAGGGGTTAAAAAAGGCC 807
```

RESULT 4
US-10-027-632-134702
; Sequence 134702, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

```
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134702
;   LENGTH: 809
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134702

  Query Match             14.5%;  Score 735.4;  DB 16;  Length 809;
  Best Local Similarity   97.3%;  Pred. No. 4.3e-205;
  Matches  789;  Conservative    1;  Mismatches   17;  Indels    4;  Gaps    4;

Qy       1312 TGCTTCCCCGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGA 1371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGCTTCCCCGTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGA 60

Qy       1372 CTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGC 1431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCGTGGGGGATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGC 120

Qy       1432 TCTCAATCGCACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCAC 1491
              |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        121 TCTCAATCGCACAGGGGACCCAGGGCCWGGCCCCTCCATCCAGAAAACCTATGACCTCAC 180

Qy       1492 CCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAG 1551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGCTACCTGGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAG 240

Qy       1552 GAATCTGGGAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGT 1611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAATCTGGGAGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGT 300

Qy       1612 TCTGGTAAATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTC 1671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTGGTAAATGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTC 360

Qy       1672 TATCTCCTGCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACT 1731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATCTCCTGCCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACT 420

Qy       1732 AGGCATGTGGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAG 1791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGCATGTGGGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAG 480

Qy       1792 CCGCCCGCCTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTC 1851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCGCCCGCCTCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTC 540

Qy       1852 TCTTGCGCGGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCA 1911
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCTTGCGCGGGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCA 600

Qy       1912 GCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCG 1971
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        601 GCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCA-GGGCCTCCCTTCCG 659

Qy       1972 TCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGG 2031
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        660 TCTGCCTCTCCCATCCTAGCTGGGCCTCCTA-GGGGGTCATGGGGGAAGGGGACTGTAGG 718

Qy       2032 GAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATT 2091
              ||||||| ||||||||||||||||| |||||||||| ||||| | |||||||||  |  |
Db        719 GAACCCA-GCAGTAGTGGCAGGGGG-TTAGGGTGTGAATGGAAGGTATGCTGTAGGATTG 776

Qy       2092 TGGGGGTGGTCCAGAGGTGTTCAGAGAGCCC 2122
              ||||| || | | ||| ||| | | || ||
Db        777 GGGGGGGGGCCAAAAGGGGTTAAAAAAGGCC 807
```

```
RESULT 5
US-10-027-632-134703
; Sequence 134703, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134703
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134703

  Query Match             12.4%;  Score 630.6;  DB 13;  Length 631;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-174;
  Matches  630;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1909 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 1968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 60

Qy       1969 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 2028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 120

Qy       2029 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 2088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 180

Qy       2089 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 2148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 240

Qy       2149 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 2208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 300

Qy       2209 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 2268
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        301 CCTACTCTGGAGCCAGGGAAAGAAAAGGGARGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 360

Qy       2269 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 2328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 420

Qy       2329 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 2388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 480

Qy       2389 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 2448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 540
```

```
Qy       2449 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 600

Qy       2509 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 2539
             |||||||||||||||||||||||||||||||
Db        601 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 631


RESULT 6
US-10-027-632-134704
; Sequence 134704, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134704
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134704

  Query Match            12.4%;  Score 630.6;  DB 13;  Length 631;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-174;
  Matches  630;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1909 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 60

Qy       1969 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 2028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 120

Qy       2029 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 180

Qy       2089 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 240

Qy       2149 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 2208
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 300

Qy       2209 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 2268
             |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db        301 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGARGGTGGCGGGGGAGCTGGCTCCAGCCC 360

Qy       2269 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 2328
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 420

Qy       2329 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 2388
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 480

Qy       2389 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 2448
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 540

Qy       2449 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 2508
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 600

Qy       2509 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 2539
             |||||||||||||||||||||||||||||||
Db        601 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 631


RESULT 7
US-10-027-632-134703
; Sequence 134703, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134703
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134703

  Query Match            12.4%;  Score 630.6;  DB 16;  Length 631;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-174;
  Matches  630;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1909 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 60

Qy       1969 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 2028
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 120

Qy       2029 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 2088
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 180

Qy       2089 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 2148
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 240

Qy       2149 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 2208
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 300

Qy       2209 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 2268
                |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        301 CCTACTCTGGAGCCAGGGAAAGAAAAGGGARGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 360

Qy       2269 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 2328
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 420

Qy       2329 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 2388
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 480

Qy       2389 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 2448
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 540

Qy       2449 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 2508
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 600

Qy       2509 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 2539
                |||||||||||||||||||||||||||||||
Db        601 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 631


RESULT 8
US-10-027-632-134704
; Sequence 134704, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 134704
;   LENGTH: 631
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-134704

  Query Match           12.4%;  Score 630.6;  DB 16;  Length 631;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-174;
  Matches  630;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1909 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 1968
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCAGCTCGCCTCTGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTT 60

Qy       1969 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 2028
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCGTCTGCCTCTCCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGT 120

Qy       2029 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 2088
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGGAACCCAGGCAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGG 180

Qy       2089 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 2148
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTGGGGGTGGTCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGG 240

Qy       2149 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 2208
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCGAGGCACCATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGAC 300

Qy       2209 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCC 2268
                |||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db        301 CCTACTCTGGAGCCAGGGAAAGAAAAGGGAAGAGRGGTGGCGGGGGAGCTGGCTCCAGCCC 360

Qy       2269 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 2328
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGGATACACCGAGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTG 420

Qy       2329 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 2388
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCCTTGCCTATCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCT 480

Qy       2389 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 2448
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCATCCCTTTCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCT 540

Qy       2449 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 2508
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTCCCTATACCATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTC 600

Qy       2509 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 2539
                |||||||||||||||||||||||||||||||
Db        601 CATTTCTTCCTTCTGCAACCCTGCTCCTATC 631


RESULT 9
US-09-931-704-1
; Sequence 1, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
;  APPLICANT: Senaldi, Giorgio
;  TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
;  TITLE OF INVENTION:  Inhibitors
;  FILE REFERENCE: A-695
;  CURRENT APPLICATION NUMBER: US/09/931,704
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: US 60/226,436
;  PRIOR FILING DATE: 2000-08-18
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 797
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (90)..(764)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (171)..()
;   OTHER INFORMATION: Met at -27
US-09-931-704-1

  Query Match           10.3%;  Score 523.4;  DB 9;  Length 797;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-142;
  Matches  524;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 330
```

```
Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 390

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 450

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 510

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 570

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 630

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 690

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 750

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 3887
             |||||||||||||||||||||||||||||||||||||||||||||
Db        751 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCC 795
```

RESULT 10
US-10-276-774-916
; Sequence 916, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 916
; LENGTH: 968
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-916

Query Match 9.7%; Score 495.4; DB 13; Length 968;
Best Local Similarity 99.8%; Pred. No. 2.2e-134;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 531

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 591

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 651

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        652 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 711

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        712 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 771

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        772 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 831

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        832 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 891

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        892 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 951

Qy       3843 GGGCTCATGGCTTCTGA 3859
             |||||||||||||||||
Db        952 GGGCTCATGGCTTCTGA 968
```

RESULT 11
US-09-864-761-23175/c
; Sequence 23175, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

```
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23175
;   LENGTH: 495
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005849.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
;   OTHER INFORMATION: SWISSPROT HIT: Q63086, EVALUE 8.00e-03
;   OTHER INFORMATION: NT HIT: gi11439486, EVALUE 0.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AI752561.1, EVALUE 0.00e+00
US-09-864-761-23175

  Query Match             9.7%;  Score 493.4;  DB 9;  Length 495;
  Best Local Similarity   99.8%;  Pred. No. 5.5e-134;
  Matches  494;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3365 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 3424
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 CTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGGCA 436

Qy       3425 GAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 3484
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 AAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAACTG 376

Qy       3485 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 3544
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 CGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCAAC 316

Qy       3545 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 3604
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 CGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCCAG 256

Qy       3605 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 3664
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 GGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGCCG 196

Qy       3665 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 3724
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 CTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGAAG 136

Qy       3725 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 3784
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 ATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGGAC 76

Qy       3785 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 3844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          75 TTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGGGG 16

Qy       3845 GCTCATGGCTTCTGA 3859
             |||||||||||||||
Db          15 GCTCATGGCTTCTGA 1


RESULT 12
US-10-296-115-607
; Sequence 607, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq Inc
;  TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
;  FILE REFERENCE: 784PCT
;  CURRENT APPLICATION NUMBER: US/10/296,115
;  CURRENT FILING DATE:  2002-11-18
;  PRIOR APPLICATION NUMBER: US09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 607
;   LENGTH: 768
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-296-115-607

  Query Match             9.7%;  Score 492.4;  DB 13;  Length 768;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-133;
  Matches  493;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 334

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 394

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 454

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 514

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 574

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 634

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        635 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 694

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        695 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 754

Qy       3843 GGGCTCATGGCTTC 3856
             ||||||||||||||
Db        755 GGGCTCATGGCTTC 768


RESULT 13
US-09-864-761-6462/c
; Sequence 6462, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
```

```
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6462
;   LENGTH: 492
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005849.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.5
US-09-864-761-6462

  Query Match             9.7%;  Score 492;  DB 9;  Length 492;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-133;
  Matches  492;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3519 TTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCC 3578
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 TTCTGTGTTACTTGCGTGGCCTCAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCC 433

Qy       3579 TGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAG 3638
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 TGGCCCACTTCTGCACCAGCCTCCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAG 373

Qy       3639 CTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCC 3698
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CTCTGGGCTACCCACTGCCCCAGCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCC 313

Qy       3699 CTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGA 3758
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 CTGCCCACAGTGACTTCCTCCAGAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGA 253

Qy       3759 CCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAG 3818
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 CCTGGCTGTGGCGCTCGGCCAAGGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAG 193

Qy       3819 CAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTT 3878
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        192 CAGCTGCAGTCACCCTGCACCTGGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTT 133

Qy       3879 CGCTCCCCCTTCAAACCCTGCTCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTG 3938
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CGCTCCCCCTTCAAACCCTGCTCCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTG 73

Qy       3939 TTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTG 3998
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TTGAGCCAGGAGACAGAAGCTGTGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTG 13

Qy       3999 ATGCGATCAGCC 4010
             ||||||||||||
Db          12 ATGCGATCAGCC 1


RESULT 14
US-09-867-701-4667
; Sequence 4667, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4667
;   LENGTH: 465
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-4667

  Query Match             8.6%;  Score 437.2;  DB 9;  Length 465;
  Best Local Similarity   98.9%;  Pred. No. 1.9e-117;
  Matches  461;  Conservative    0;  Mismatches    3;  Indels    2;  Gaps    2;

Qy        133 CAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGCGCCCCATCTGA 192
              || | |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db          1 CAAGCTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAGC-CCCCATCTGA 59

Qy        193 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 TACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCCCAGACCTCACCA 119

Qy        253 CATCCTTG-TGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 311
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 CATCCTTGCTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCAAGTCTAAACTA 179

Qy        312 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 371
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 ATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAATGCCAATTATAT 239

Qy        372 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        240 CTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTGAATCCTCACCT 299

Qy        432 TGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 491
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TGGTCCTCACTTAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACATCCCTCCTTACA 359

Qy        492 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 GCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACTCTGCAGTGCCA 419

Qy        552 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTG 597
              |||||||||||||||||||||||||||||||||||||||||||||
Db        420 TCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTG 465
```

```
RESULT 15
US-09-931-704-4
; Sequence 4, Application US/09931704
; Patent No. US20020041873A1
; GENERAL INFORMATION:
;  APPLICANT: Senaldi, Giorgio
;  TITLE OF INVENTION: Methods and Compositions for Treating IgE-Related Disease Using N
;  TITLE OF INVENTION:  Inhibitors
;  FILE REFERENCE: A-695
;  CURRENT APPLICATION NUMBER: US/09/931,704
;  CURRENT FILING DATE:  2001-08-16
;  PRIOR APPLICATION NUMBER: US 60/226,436
;  PRIOR FILING DATE: 2000-08-18
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 819
;   TYPE: DNA
;   ORGANISM: Murine
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (95)..(769)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (176)..()
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptid
;   LOCATION: (176)..(769)
;   OTHER INFORMATION:
;   NAME/KEY: sig_peptide
;   LOCATION: (95)..(175)
;   OTHER INFORMATION:
US-09-931-704-4

  Query Match             8.4%;  Score 429.6;  DB 9;  Length 819;
  Best Local Similarity   90.4%;  Pred. No. 4.6e-115;
  Matches  459;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||| |||||||||||||||||| |||||||| ||||| || |||||||
Db        276 ACCTGAACTACCTGGGGCCCCCTTTCAACGAGCCTGACTTCAATCCTCCTCGACTGGGGG 335

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             |||| |||||||||||||||| ||  ||||||| |||||||||||||||||||||| |
Db        336 CAGAAACTCTGCCCAGGGCCACGGTCAACTTGGAAGTGTGGCGAAGCCTCAATGACAGGC 395

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             |||||||||||||||||||| ||||| ||||| ||||| ||||||||||||||||||||||
Db        396 TGCGGCTGACCCAGAACTATGAGGCGTACAGTCACCTCCTGTGTTACTTGCGTGGCCTCA 455

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             |||||||||||||||| ||||| || || || |||||||||||||||| ||||||||||
Db        456 ACCGTCAGGCTGCCACAGCTGAACTCCGACGTAGCCTGGCCCACTTCTGTACCAGCCTCC 515

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             |||||||||||||||||||| || ||||||||  | || ||||||||||||||||||||
Db        516 AGGGCCTGCTGGGCAGCATTGCAGGTGTCATGGCGACGCTTGGCTACCCACTGCCCCAGC 575

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             | ||||| |||||||| ||  | ||| | ||||||||||||||||||||||||||||||||
Db        576 CTCTGCCAGGGACTGAGCCAGCCTGGGCCCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 635

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||| |||||||||||||||||||||||||||||||||| ||||| || |||||||
Db        636 AGATGGATGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTATGGCGTTCAGCCAAGG 695

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             |||||||||||| |||||||||||||||||||||||||| |||||||||||||| |||
Db        696 ACTTCAACCGGCTTAAGAAGAAGATGCAGCCTCCAGCAGCTTCAGTCACCCTGCACTTGG 755

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTT 3870
                |||  ||||| ||||||| ||||||| |
Db        756 AGGCACATGGTTTCTGACCTCTGACCCT 783
```

Search completed: August 16, 2004, 10:40:07
Job time : 2944.55 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 16:10:45 ; Search time 374.903 Seconds
(without alignments)
7530.032 Million cell updates/sec

Title: US-09-931-704-3
Perfect score: 5087
Sequence: 1 aacctgcgagtgggcctggc.....cctttgctaagtcttcctca 5087

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	5087	100.0	5087	1	US-08-792-019B-3	Sequence 3, Appli
	2	5087	100.0	5087	3	US-08-988-819-3	Sequence 3, Appli
	3	5087	100.0	5087	3	US-09-016-534-3	Sequence 3, Appli
	4	1463.4	28.8	1710	3	US-09-106-182-1	Sequence 1, Appli
	5	523.4	10.3	797	1	US-08-792-019B-1	Sequence 1, Appli
	6	523.4	10.3	797	3	US-08-988-819-1	Sequence 1, Appli
	7	523.4	10.3	797	3	US-09-016-534-1	Sequence 1, Appli
	8	429.6	8.4	819	1	US-08-792-019B-4	Sequence 4, Appli
	9	429.6	8.4	819	3	US-08-988-819-4	Sequence 4, Appli
	10	429.6	8.4	819	3	US-09-016-534-4	Sequence 4, Appli
	11	324.6	6.4	396	3	US-09-106-182-7	Sequence 7, Appli
	12	99	1.9	7218	1	US-08-232-463-14	Sequence 14, Appl
	13	60.2	1.2	1926	4	US-09-249-585A-4	Sequence 4, Appli
	14	60.2	1.2	1931	2	US-09-130-114-2	Sequence 2, Appli
	15	57.6	1.1	53526	3	US-08-658-136-2	Sequence 2, Appli
	16	57.6	1.1	53577	3	US-08-658-136-1	Sequence 1, Appli
	17	52.4	1.0	51259	3	US-08-781-891-209	Sequence 209, App
	18	52.4	1.0	51259	4	US-09-618-166-209	Sequence 209, App
	19	52.2	1.0	1166	4	US-09-072-596-323	Sequence 323, App
	20	52.2	1.0	1166	4	US-09-072-967-328	Sequence 328, App
	21	49.6	1.0	16442	3	US-08-781-891-208	Sequence 208, App
	22	49.6	1.0	16442	4	US-09-618-166-208	Sequence 208, App
c	23	48.8	1.0	1926	4	US-09-249-585A-2	Sequence 2, Appli
c	24	48.8	1.0	1926	4	US-09-410-399-3	Sequence 3, Appli
c	25	48.8	1.0	2580	3	US-09-050-863-2	Sequence 2, Appli
c	26	48.8	1.0	2580	4	US-09-359-081-2	Sequence 2, Appli
	27	48.8	1.0	5452	2	US-09-130-114-1	Sequence 1, Appli
c	28	48.8	1.0	7218	1	US-08-232-463-14	Sequence 14, Appl
	29	48.8	1.0	8705	4	US-09-647-344A-14	Sequence 14, Appl
c	30	48.8	1.0	9600	3	US-08-910-647-1	Sequence 1, Appli
c	31	48.8	1.0	9600	4	US-09-620-925-1	Sequence 1, Appli
c	32	48.8	1.0	10596	1	US-07-884-811-15	Sequence 15, Appl
c	33	48.8	1.0	10596	1	US-07-885-971-15	Sequence 15, Appl
c	34	48.8	1.0	10596	1	US-08-087-783A-15	Sequence 15, Appl
c	35	48.8	1.0	10596	1	US-08-194-088B-15	Sequence 15, Appl
c	36	48.8	1.0	10596	2	US-08-194-087-15	Sequence 15, Appl
c	37	48.8	1.0	10596	5	PCT-US93-04648-15	Sequence 15, Appl
	38	48.8	1.0	16080	4	US-09-724-566A-48	Sequence 48, Appl
c	39	48.6	1.0	289	3	US-09-007-005-17	Sequence 17, Appl
c	40	48.6	1.0	289	3	US-09-244-796-17	Sequence 17, Appl
c	41	48	0.9	1866	3	US-09-173-581-13	Sequence 13, Appl
c	42	48	0.9	1866	3	US-09-420-915-13	Sequence 13, Appl
	43	45.2	0.9	950	3	US-08-991-789A-6	Sequence 6, Appli
	44	45.2	0.9	950	4	US-09-062-451-6	Sequence 6, Appli
	45	45.2	0.9	950	4	US-09-598-326-6	Sequence 6, Appli

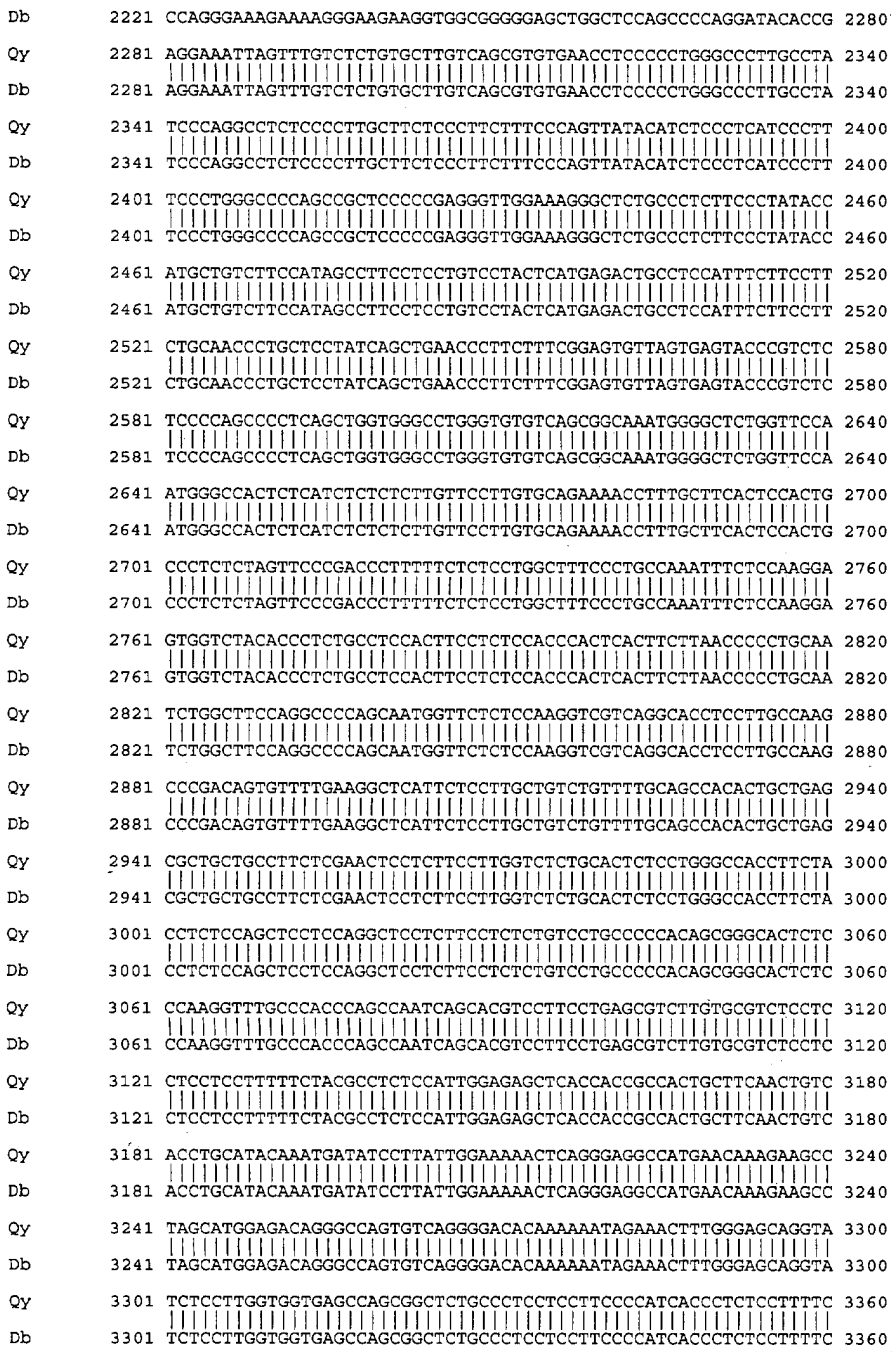
ALIGNMENTS

RESULT 1
US-08-792-019B-3
; Sequence 3, Application US/08792019B
; Patent No. 5741772
; GENERAL INFORMATION:
; APPLICANT: CHANG, MING-SHI
; TITLE OF INVENTION: THE NEUROTROPHIC FACTOR NNT-1
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMGEN INC.
; STREET: 1840 DEHAVILLAND DRIVE
; CITY: THOUSAND OAKS
; STATE: CA
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,019B
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: COOK, ROBERT R.
; REGISTRATION NUMBER: 31,602
; REFERENCE/DOCKET NUMBER: A-442
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5087 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 137..138
; OTHER INFORMATION: /product= "INTERVENING UNSEQUENCED
; OTHER INFORMATION: REGION OF >1KB"
US-08-792-019B-3

Query Match 100.0%; Score 5087; DB 1; Length 5087;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5087; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
```


Qy 61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
Db 61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
Qy 121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Db 121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
Qy 181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Db 181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
Qy 241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
Db 241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
Qy 301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
Db 301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
Qy 361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
Db 361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
Qy 421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
Db 421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
Qy 481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
Db 481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
Qy 541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
Db 541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
Qy 601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
Db 601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
Qy 661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
Db 661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
Qy 721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
Db 721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
Qy 781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
Db 781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
Qy 841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Db 841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
Qy 901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Db 901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
Qy 961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Db 961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
Qy 1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Db 1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
Qy 1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Db 1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
Qy 1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Db 1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
Qy 1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Db 1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
Qy 1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Db 1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
Qy 1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Db 1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
Qy 1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Db 1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
Qy 1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Db 1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Qy 1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Db 1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Qy 1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Db 1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Qy 1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Db 1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Qy 1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Db 1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Qy 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db 1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Qy 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db 1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Qy 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db 1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Qy 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db 1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Qy 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db 1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db 2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db 2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db 2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280


Db 2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db 2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db 2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db 2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db 2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db 2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db 2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db 2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db 2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db 2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db 2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db 2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db 2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db 3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db 3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db 3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db 3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db 3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db 3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db 3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db 3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db 3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db 3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db 3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Db 3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Db 3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Db 3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Db 3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Db 3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Db 3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Db 4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
Qy 4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Db 4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy 4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Db 4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy 4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Db 4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Qy 4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Db 4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Qy 4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Db 4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Qy 4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Db 4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440

```
Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500

Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

```
RESULT 2
US-08-988-819-3
; Sequence 3, Application US/08988819
; Patent No. 6054294
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/988,819
;      FILING DATE:  12-DEC-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442A
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5087 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  137..138
;      OTHER INFORMATION:  /product= "INTERVENING UNSEQUENCED
;      OTHER INFORMATION:  REGION OF >1KB"
US-08-988-819-3

  Query Match             100.0%;  Score 5087;  DB 3;  Length 5087;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
```

```
Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900

Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960

Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020

Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080

Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140

Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200

Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260

Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320

Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380

Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440

Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500

Qy       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560

Qy       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620

Qy       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680

Qy       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740

Qy       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800

Qy       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860

Qy       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920

Qy       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980

Qy       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040

Qy       2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100

Qy       2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160

Qy       2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220

Qy       2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280

Qy       2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340

Qy       2341 TCCCAGGCCTCTCCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCCCAGGCCTCTCCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400

Qy       2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460

Qy       2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520

Qy       2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580

Qy       2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640

Qy       2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700

Qy       2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760

Qy       2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820

Qy       2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880

Qy       2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940

Qy       2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy       3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy       3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy       3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy       3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy       3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy       3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy       3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy       3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy       3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy       3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy       3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Qy       3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
Qy       3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
Qy       3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
Qy       3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
Qy       3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
Qy       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
Qy       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
```

```
Qy       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
Qy       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
Qy       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
Qy       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
Qy       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 3
US-09-016-534-3
; Sequence 3, Application US/09016534

```
; Patent No. 6143874
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    APPLICANT:  ELLIOTT, GARY S.
;    APPLICANT:  SARMIENTO, ULLA
;    APPLICANT:  SENALDI, GIORGIO
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  ONE AMGEN CENTER
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,534
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/792,019
;      FILING DATE:  03-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442B
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5087 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  137..138
;      OTHER INFORMATION:  /product= "INTERVENING UNSEQUENCED
;      OTHER INFORMATION:  REGION OF >1KB"
US-09-016-534-3

  Query Match             100.0%;  Score 5087;  DB 3;  Length 5087;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 5087;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AACCTGCGAGTGGGCCTGGCGGATGGGATTATTAAAGCTTCGCCGGAGCCGCGGCTCGCC 60

Qy         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCCCACTCCGCCAGCCTCCGGGAGAGGAGCCGCACCCGGCCGGCCCAGCCCCAGCCCCA 120

Qy        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGGACCTCCGAGCAGGTTGAAAACCCAAACTAGCCCTGCTCTTCATAACATGACAAGCAG 180

Qy        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGCCCCATCTGATACCTAAACCGACCAAGTCACAGCCCTCCAACTCACCCTCTGCCTGCC 240

Qy        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CAGACCTCACCACATCCTTGTGGACTCAAACCTCAACCGCACTAAATCAACCAAATCCCA 300

Qy        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGTCTAAACTAATCTGAAACTTTTAAAGTAACCCAGTCCTTAAACCTAACCTAGCCCAAT 360

Qy        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCAATTATATCTACCCTAGCCAAACCCTAACTGCCTTTGCCAGTCCAAAGTGTCCACTG 420

Qy        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATCCTCACCTTGGTCCTCACTGAAAATCCCAGAAAAGCATATTTCCCCACTGCCCACAT 480

Qy        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCCTCCTTACAGCACCCAACCCTGGCCTCTGGACTCCTGGTATCCTGGGATGTCCAAACT 540

Qy        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTGCAGTGCCATCAGCCAACAAGCCCGACTCGTCAAATGCACCTCTCTCCCTTCCTGTCC 600

Qy        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCACCCTTGCAGGCTGATGGAAAGGCCTCATTGAAGTCCAACTTTTCCCCACCTAACACC 660

Qy        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGAACGGGGTGAACCTCCACACTGCCACCGTTCCCTGAGAGTGAGCACTAAATCTCCTT 720

Qy        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAATCTAACCCCACCCTACACTTCCCACACTCAGGAATCACATCCTAGAATATACCCAAA 780

Qy        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACTAAGCCCCATAAGGCAGCCCGACCCTAGTGGTCTAACCCTATACCTTGCTTCCTATGG 840

Qy        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGAGTCTGTTCTTGGCGGCCGCCTCTCTCCTGCTTCCTCCCTTAGAGCTGACTGTGCTC 900

Qy        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGCCTGCCAGCTCTGACATGTGCTGTCTCCCACCCTCTGACTCCCCTCAAGCTGCAGTGG 960

Qy        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GACTGGAAGACTGGCAGGAAGCTAGGGTACAACTGGAACACAGGCAGGTCGACCTGCAGT 1020

Qy       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCCTAGGCCTGGCCCCGTCCCTCCATGTACACACATATACATGTTGGCACACACACAGTG 1080

Qy       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GCACACATGCCAAAGACTCTCTCAGCTGACACACAGATCCATTCTCAAGTATCTACTGAT 1140

Qy       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AGACACTCATGCGTGCCAAGTCCTCATCCTCAAACATACACATGCCTCTCTTTCTCTCCC 1200

Qy       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCTTGCCAGGAGTGTTTCCCCTCCTCCATCCCCTCTGCCTCCCATCTGGTGTCCCACCC 1260

Qy       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TCACCCCCCACCCAGCCCAAGGTGGGGACAGACACCTGAGGGGCTGCCAGCTGCTTCCCC 1320

Qy       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGTGGGCCCGGGCCGCGCTCATGCTTCTCGTCCATCCTGCCCACAGGGGACTCGTGGGG 1380

Qy       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GATGTTAGCGTGCCTGTGCACGGTGCTCTGGCACCTCCCTGCAGTGCCAGCTCTCAATCG 1440

Qy       1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
```

```
Db      1441 CACAGGGGACCCAGGGCCTGGCCCCTCCATCCAGAAAACCTATGACCTCACCCGCTACCT 1500
Qy      1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Db      1501 GGAGCACCAACTCCGCAGCTTGGCTGGGACCTATGTGAGTATCCAGCGTAGGAATCTGGG 1560
Qy      1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Db      1561 AGTTGGGGAGGAGTGAGGAGTTGGGGAAAGACAGTCCTAACCGTGGAGGGTTCTGGTAAA 1620
Qy      1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Db      1621 TGATGGGGTGAGGAGGGGCTCTTTGGCTCCCACCAGTCCCCCTGTCTGGTCTATCTCCTG 1680
Qy      1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Db      1681 CCCTTCCCTCTTAGGTGGCCCCCCCACTTCCCCATCCCTGGCCCCAGGACTAGGCATGTG 1740
Qy      1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Db      1741 GGCAGGCCTCGCACCCGCCTTGGCCCATTGCCCCACTGGCTGCCAGCCCAGCCGCCCGCC 1800
Qy      1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Db      1801 TCCCCCTGGGGGCCGGGGAAGTCTCCTCTGTTTACACCGTGTTGTGGTGTCTCTTGCGCG 1860
Qy      1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Db      1861 GGCGGGGTTGGGTGGGGACAGAGGGGCCCCACCTCCCATGCCTGCGTTCCAGCTCGCCTC 1920
Qy      1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Db      1921 TGCCCCCAGACCTGGGGCCCTGCTGCTCTGGACCCAGGGGCCTCCCTTCCGTCTGCCTCT 1980
Qy      1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Db      1981 CCCATCCTAGCTGGGCCTCCTAGGGGGGTCATGGGGGAAGGGGACTGTAGGGAACCCAGG 2040
Qy      2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Db      2041 CAGTAGTGGCAGGGGGTTTAGGGTGTGGATGGAGGTTATGCTGTAAGGATTTGGGGGTGG 2100
Qy      2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Db      2101 TCCAGAGGTGTTCAGAGAGCCCAGGAGAGAAGGAAGGAGGGTTGGAGGAGCCGAGGCACC 2160
Qy      2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Db      2161 ATGGGGAACCGGCCCCCTCTTCCCGTGTTCCTCTTCCACATCCCAGACCCTACTCTGGAG 2220
Qy      2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Db      2221 CCAGGGAAAGAAAAGGGAAGAAGGTGGCGGGGGAGCTGGCTCCAGCCCCAGGATACACCG 2280
Qy      2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Db      2281 AGGAAATTAGTTTGTCTCTGTGCTTGTCAGCGTGTGAACCTCCCCCTGGGCCCTTGCCTA 2340
Qy      2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Db      2341 TCCCAGGCCTCTCCCCTTGCTTCTCCCTTCTTTCCCAGTTATACATCTCCCTCATCCCTT 2400
Qy      2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Db      2401 TCCCTGGGCCCCAGCCGCTCCCCCGAGGGTTGGAAAGGGCTCTGCCCTCTTCCCTATACC 2460
Qy      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Db      2461 ATGCTGTCTTCCATAGCCTTCCTCCTGTCCTACTCATGAGACTGCCTCCATTTCTTCCTT 2520
Qy      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Db      2521 CTGCAACCCTGCTCCTATCAGCTGAACCCTTCTTTCGGAGTGTTAGTGAGTACCCGTCTC 2580
Qy      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Db      2581 TCCCCAGCCCCTCAGCTGGTGGGCCTGGGTGTGTCAGCGGCAAATGGGGCTCTGGTTCCA 2640
Qy      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Db      2641 ATGGGCCACTCTCATCTCTCTCTTGTTCCTTGTGCAGAAAACCTTTGCTTCACTCCACTG 2700
Qy      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Db      2701 CCCTCTCTAGTTCCCGACCCTTTTTCTCTCCTGGCTTTCCCTGCCAAATTTCTCCAAGGA 2760
Qy      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Db      2761 GTGGTCTACACCCTCTGCCTCCACTTCCTCTCCACCCACTCACTTCTTAACCCCCTGCAA 2820
Qy      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Db      2821 TCTGGCTTCCAGGCCCCAGCAATGGTTCTCTCCAAGGTCGTCAGGCACCTCCTTGCCAAG 2880
Qy      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Db      2881 CCCGACAGTGTTTTGAAGGCTCATTCTCCTTGCTGTCTGTTTTGCAGCCACACTGCTGAG 2940
Qy      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Db      2941 CGCTGCTGCCTTCTCGAACTCCTCTTCCTTGGTCTCTGCACTCTCCTGGGCCACCTTCTA 3000
Qy      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Db      3001 CCTCTCCAGCTCCTCCAGGCTCCTCTTCCTCTCTGTCCTGCCCCCACAGCGGGCACTCTC 3060
Qy      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Db      3061 CCAAGGTTTGCCCACCCAGCCAATCAGCACGTCCTTCCTGAGCGTCTTGTGCGTCTCCTC 3120
Qy      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Db      3121 CTCCTCCTTTTTCTACGCCTCTCCATTGGAGAGCTCACCACCGCCACTGCTTCAACTGTC 3180
Qy      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Db      3181 ACCTGCATACAAATGATATCCTTATTGGAAAAACTCAGGGAGGCCATGAACAAAGAAGCC 3240
Qy      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Db      3241 TAGCATGGAGACAGGGCCAGTGTCAGGGGACACAAAAAATAGAAACTTTGGGAGCAGGTA 3300
Qy      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Db      3301 TCTCCTTGGTGGTGAGCCAGCGGCTCTGCCCTCCTCCTTCCCCATCACCCTCTCCTTTTC 3360
Qy      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Db      3361 ACAGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGG 3420
Qy      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Db      3421 GGCAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAA 3480
Qy      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Db      3481 ACTGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCT 3540
Qy      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Db      3541 CAACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCT 3600
Qy      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
Db      3601 CCAGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCA 3660
```

```
Qy       3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GCCGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCA 3720

Qy       3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAAGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAA 3780

Qy       3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 GGACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCT 3840

Qy       3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 GGGGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCT 3900

Qy       3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 CCCACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTG 3960

Qy       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TGAGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTC 4020

Qy       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 CCCACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTG 4080

Qy       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 TTTCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCT 4140

Qy       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 CATGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCA 4200

Qy       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CATCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCT 4260

Qy       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TTGGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACA 4320

Qy       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 TCCAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTG 4380

Qy       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 CAGGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCAC 4440

Qy       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 AGGGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAA 4500

Qy       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACCCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTC 4560

Qy       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCTGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCAC 4620

Qy       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 ACCACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCT 4680

Qy       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 CCCCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATAT 4740

Qy       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TTATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTA 4800

Qy       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 TTAACAAATAAAATGCTTGCCCCAGAACTTAGTCTCTTTGCCCAGCCTCACCCCTCCTGG 4860

Qy       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 TGCTCATCAGACTCTTGCCACCCCTGGCTCCCACTCCCTGCTTGCCTCTGGTGGAGCTGC 4920

Qy       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ACAGAGCTCTGGGAAGAGGCCCTCTTCCTCCCCGCACTGGGGCGATGGGCGCACCTCAGA 4980

Qy       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTTACCCACTGCTGCTGCCACCACCAACCCCTTGATCCCTCAGTCCTCCCACACAGCTTC 5040

Qy       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 TGTCCACCCCAGGTTTCCCTCACCCCACCTTTGCTAAGTCTTCCTCA 5087
```

RESULT 4
US-09-106-182-1
; Sequence 1, Application US/09106182
; Patent No. 6046035
; GENERAL INFORMATION:
; APPLICANT: Shi, Yanggu
; APPLICANT: Ruben, Steve
; TITLE OF INVENTION: Cardiotrophin-Like Cytokine
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc
; STREET: 9410 Key West Ave
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,182
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/051,053
; FILING DATE: 30-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF385
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 46..720
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 46..126

```
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  127..720
US-09-106-182-1

  Query Match             28.8%;  Score 1463.4;  DB 3;  Length 1710;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1464;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3363 AGCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 3422
             | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATCTGAACTACCTGGGCCCCCCTTTCAACGAGCCAGACTTCAACCCTCCCCGCCTGGGGG 286

Qy       3423 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 3482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 CAGAGACTCTGCCCAGGGCCACTGTTGACTTGGAGGTGTGGCGAAGCCTCAATGACAAAC 346

Qy       3483 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 3542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TGCGGCTGACCCAGAACTACGAGGCCTACAGCCACCTTCTGTGTTACTTGCGTGGCCTCA 406

Qy       3543 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 3602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 ACCGTCAGGCTGCCACTGCTGAGCTGCGCCGCAGCCTGGCCCACTTCTGCACCAGCCTCC 466

Qy       3603 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 3662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGGGCCTGCTGGGCAGCATTGCGGGCGTCATGGCAGCTCTGGGCTACCCACTGCCCCAGC 526

Qy       3663 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 3722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CGCTGCCTGGGACTGAACCCACTTGGACTCCTGGCCCTGCCCACAGTGACTTCCTCCAGA 586

Qy       3723 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 3782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AGATGGACGACTTCTGGCTGCTGAAGGAGCTGCAGACCTGGCTGTGGCGCTCGGCCAAGG 646

Qy       3783 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 3842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 ACTTCAACCGGCTCAAGAAGAAGATGCAGCCTCCAGCAGCTGCAGTCACCCTGCACCTGG 706

Qy       3843 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GGGCTCATGGCTTCTGACTTCTGACCTTCTCCTCTTCGCTCCCCCTTCAAACCCTGCTCC 766

Qy       3903 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CACTTTGTGAGAGCCAGCCCTGTATGCCAACACCTGTTGAGCCAGGAGACAGAAGCTGTG 826

Qy       3963 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 4022
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 AGCCTCTGGCCCTTTCCTGGACCGGCTGGGCGTGTGATGCGATCAGCCCTGTCTCCTCCC 886

Qy       4023 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 4082
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CACCTCCCAAAGGTCTACCGAGCTGGGGAGGAGGTACAGTAGGCCCTGTCCTGTCCTGTT 946

Qy       4083 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 4142
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 TCTACAGGAAGTCATGCTCGAGGGAGTGTGAAGTGGTTCAGGTTGGTGCAGAGGCGCTCA 1006

Qy       4143 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 4202
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TGGCCTCCTGCTTCTTGCCTACCACTTGGCCAGTGCCCACCCAGCCCCTCAGGTGGCACA 1066

Qy       4203 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 4262
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 TCTGGAGGGCAGGGGTTGAGGGGCCACCACCACACATGCCTTTCTGGGGTGAAGCCCTTT 1126

Qy       4263 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 4322
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 GGCTGCCCCACTCTCCTTGGATGGGTGTTGCTCCCTTATCCCCAAATCACTCTATACATC 1186

Qy       4323 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 4382
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAATTCAGGAAACAAACATGGTGGCAATTCTACACAAAAAGAGATGAGATTAACAGTGCA 1246

Qy       4383 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 4442
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGGTTGGGGTCTGCATTGGAGGTGCCCTATAAACCAGAAGAGAAAATACTGAAAGCACAG 1306

Qy       4443 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 4502
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 GGGCAGGGACAGACCAGACCAGACCCAGGAGTCTCCAAAGCACAGAGTGGCAAACAAAAC 1366

Qy       4503 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 4562
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CCGAGCTGAGCATCAGGACCTTGCCTCGAATTGTCTTCCAGTATTACGGTGCCTCTTCTC 1426

Qy       4563 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 4622
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1427 TGCCCCCTTTCCCAGGGTATCTGTGGGTTGCCAGGCTGGGGAGGGCAACCATAGCCACAC 1486

Qy       4623 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 4682
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1487 CACAGGATTTCCTGAAAGTTTACAATGCAGTAGCATTTTGGGGTGTAGGGTGGCAGCTCC 1546

Qy       4683 CCAAGGCCCTGCCCCCCAGCCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 4742
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1547 CCAAGGCCCTGCCCCCCAGCCCCACCCACTCATGACTCTAAGTGTGTTGTATTAATATTT 1606

Qy       4743 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 4802
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1607 ATTTATTTGGAGATGTTATTTATTAGATGATATTTATTGCAGAATTTCTATTCTTGTATT 1666

Qy       4803 AACAAATAAAATGCTTGCCCCAGAA 4827
             |||||||||||||||||||||||||
Db       1667 AACAAATAAAATGCTTGCCCCAGAA 1691


RESULT 5
US-08-792-019B-1
; Sequence 1, Application US/08792019B
; Patent No. 5741772
;  GENERAL INFORMATION:
;    APPLICANT:  CHANG, MING-SHI
;    TITLE OF INVENTION:  THE NEUROTROPHIC FACTOR NNT-1
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AMGEN INC.
;      STREET:  1840 DEHAVILLAND DRIVE
;      CITY:  THOUSAND OAKS
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/792,019B
;      FILING DATE:  03-FEB-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOK, ROBERT R.
;      REGISTRATION NUMBER:  31,602
;      REFERENCE/DOCKET NUMBER:  A-442
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  797 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
```